Effects of Educational Investments Within and Beyond a Framework of Home-Based Parental Involvement. Qualitative and Quantitative Analyses on South African Subpopulations.

Doctoral Thesis
for the Attainment of the Degree of
Doctor of Philosophy (Dr. phil.)
of the Faculty of Education at the University of Hamburg

Submitted by
Nadine Twele

Hamburg, September 02, 2022

1st supervisor & chair of the examination board:
Prof. Dr. Knut Schwippert
University of Hamburg/Germany, Faculty of Education, EW 1

2nd supervisor:
Prof. Dr. Ann-Katrin van den Ham
University of Hamburg/Germany, Faculty of Education, EW 1

3rd supervisor:
Prof. Surette van Staden, Ph.D.
University of Pretoria/South Africa, Faculty of Education, CEA

Date of Defense of the Doctoral Thesis:
January 13, 2023

Acknowledgments:

To the DVV International Project Office South Africa and The Big Issue South Africa for supporting and facilitating major parts of this thesis.

To Prof. Dr. Knut Schwippert, Prof. Dr. Ann-Katrin van den Ham, and Prof. Surette van Staden, Ph.D. for guiding, supervising and assessing this thesis.

To my former colleagues at IEA Hamburg for their support and offering in expertise.

To Justin Wild, Ph.D. for his professional advice and editorial comments.

"Because there is nothing we can do without education."

Street Vendor in Cape Town/South Africa, 2013

Abstract (English)

Educational investments are generally known to enhance life opportunities on a global scale. There are various approaches to educational investment, but they can be thought of broadly as actions that support educational attainment. Thus, they can be considered a lever for upward mobility. However, based on theories of rational decision making, it is assumed that educational investment behavior is highly affected by the actors' socioeconomic status. That is, individuals of lower socioeconomic status are less likely to make educational investments due to cost-benefit calculations. Such calculations are a phenomenon of Expectancy-Value Theory and indicate that actional behavior relies on two factors, namely costs and benefits of an action alternative. Corresponding rational decision making is often observed among parents with respect to educational investments for their children's educational trajectories. It is assumed, accordingly, that parental educational investment behavior has a major effect on children's life opportunities, and thus, influences the inheritance of socioeconomic status. This chain of events is particularly detrimental for children of lower socioeconomic background. Such effects have been primarily investigated in the U.S. and European countries. Seeking to understand the mechanisms of educational investment behavior of those with lower socioeconomic backgrounds (referred to as "socioeconomic low-performers" in this study) in highly segregated societies such as South Africa, this doctoral thesis addresses three key questions: (I) are socioeconomic low-performers in South Africa motivated to invest in education; (II) is the motivation transferred into (parental) educational investments; and (III) are parental educational investments beneficial to student outcomes in learning.

In order to answer these research questions, a qualitative and a quantitative study is conducted, each based on a separate sample and addressing educational investment behavior from a different angle. First, a qualitative approach is utilized to evaluate whether motivation for educational investments is evident among socioeconomic low-performers, and if this motivation is selectively put into action. Qualitative analyses are based on a convenience sample of South African adults (n = 12), and claims to be exploratory

with regard to general attitudes of educational investment behavior. Second, quantitative analyses are applied to investigate a specific form of educational investment behavior, namely parental involvement activities, and explore whether those activities are positively associated with student outcomes in learning. The quantitative approach is based on a nationally representative sample of South African fourth-grade students (n = 15,744) collected in prePIRLS 2011. Both empirical approaches are oriented along the lines of a framework based on psychological and sociological theories of human action behavior, parental involvement, and rational decision making, distinguishing between phases of motivation, volition and evaluation.

Results of the qualitative study show that socioeconomic low-performers in South Africa are motivated overall to invest in education, and occasionally put respective motivation into actions. Reasons for educational non-investments include the lack of time and monetary means. With regard to parental educational investment behavior, results of the quantitative study show that socioeconomic low-performing parents are generally motivated to invest in their children's education as well. Thus, they are highly involved. However, in comparison to socioeconomic high-performers, parental involvement among socioeconomic low-performs is not significantly associated with student outcomes in learning.

To summarize, results of this study reveal that the main obstacle is not (parental) motivation for educational investments or the making of educational investments, but rather with the quality thereof. With regard to the Apartheid Regime and its long-term consequences, especially for socioeconomic low-performers in South Africa, the salient question is whether the enhancement of the quality of parental educational investments holds the potential to make a difference in children's educational performance and careers, respectively; and thus, to help mitigate the inheritance of social deprivation. Results of this study provide evidence for the assumption that this is true, and thus, encourage further analyses on the subject.

Abstract (German)

Es ist allgemein bekannt, dass Bildungsinvestitionen Lebenschancen verbessern können. Es gibt verschiedene Ansätze der Definition von Bildungsinvestitionen, allgemein werden sie aber als Maßnahmen betrachtet, die das Erreichen von Bildungszielen unterstützen. Sie werden daher auch als Hebel für Aufwärtsmobilität angesehen. Auf der Grundlage von Theorien zur rationalen Entscheidungsfindung wird davon ausgegangen, dass das Bildungsinvestitionsverhalten in hohem Maße von den sozioökonomischen Ressourcen der Akteure abhängig ist. Es wird angenommen, dass Bildungsinvestitionen unter Personen mit einem niedrigeren sozioökonomischen Status aufgrund von Kosten-Nutzen Rechnungen weniger wahrscheinlich sind. Entsprechende Abwägungen sind ein Phänomen der Wert-Erwartungstheorie und zeigen, dass das Handlungsverhalten von zwei Faktoren abhängt, nämlich von den Kosten und dem Nutzen einer Handlungsalternative. Eine entsprechend rationale Entscheidungsfindung ist häufig bei Eltern zu beobachten, und zwar in Bezug auf Bildungsinvestitionen für die Bildungskarriere ihrer Kinder. Folglich wird davon ausgegangen, dass das elterliche Bildungsinvestitionsverhalten die Lebenschancen von Kindern maßgeblich mitbestimmt und somit die Vererbung des sozioökonomischen Status beeinflusst. Diese Verkettung von Ereignissen wirkt sich besonders nachteilig für Kinder mit niedrigerem sozioökonomischem Hintergrund aus. Entsprechende Effekte sind vor allem für U.S. amerikanische und europäische Länder untersucht worden. Um selbige Mechanismen des Bildungsinvestitionsverhaltens von sozioökonomisch benachteiligten Personen in hochgradig segregierten Gesellschaften wie Südafrika zu verstehen, werden in dieser Dissertation drei Schlüsselfragen untersucht: (I) Sind sozioökonomisch benachteiligte Personen in Südafrika motiviert, in Bildung zu investieren; (II) wird diese Motivation in Bildungsinvestitionen, auch im Sinne elterlicher Aktivitäten, umgesetzt; und (III) wirken sich elterliche Bildungsinvestitionen positiv auf den Bildungserfolg von Schüler:innen aus.

Zur Beantwortung dieser Forschungsfragen werden eine qualitative und eine quantitative Studie durchgeführt, die jeweils auf einer separaten Stichprobe basieren und das Bildungsinvestitionsverhalten aus unterschiedlichen Perspektiven betrachten. Erstens wird

ein qualitativer Ansatz verwendet, um zu evaluieren, ob sozioökonomisch benachteiligte Personen Motivation für Bildungsinvestitionen zeigen und ob diese Motivation selektiv in die Tat umgesetzt wird. Die qualitative Analyse basiert auf einer Stichprobe südafrikanischer Straßenhändler:innen (n = 12) und erhebt den Anspruch, allgemeine Einstellungen zum Bildungsinvestitionsverhalten zu erforschen. Zweitens werden quantitative Analysen durchgeführt, um eine spezifische Form des Bildungsinvestitionsverhaltens zu untersuchen, nämlich elterliche Aktivitäten, und ob diese Aktivitäten positiv mit den Lernergebnissen von Schüler:innen verbunden sind. Der quantitative Ansatz basiert auf einer landesweit repräsentativen Stichprobe südafrikanischer Viertklässler:innen (n = 15.744), die im Rahmen von prePIRLS 2011 erhoben wurde. Beide empirischen Ansätze orientieren sich an einem konzeptionellen Rahmen, der auf psychologischen und soziologischen Theorien des menschlichen Handlungsverhaltens, des elterlichen Engagements und der rationalen Entscheidungsfindung beruht, wobei zwischen Phasen der Motivation, der Volition und der Evaluation unterschieden wird.

Die Ergebnisse der qualitativen Studie deuten darauf hin, dass sozioökonomisch benachteiligte Personen in Südafrika insgesamt motiviert sind, in Bildung zu investieren, und dass entsprechende Motivation vereinzelt in Bildungsinvestitionen resultiert. Gründe für das Ausbleiben von Bildungsinvestitionen sind der Mangel von Zeit und finanziellen Mitteln. In Bezug auf das elterliche Bildungsinvestitionsverhalten zeigen die Ergebnisse der quantitativen Studie, dass sozioökonomisch benachteiligte Eltern generell motiviert sind, in die Bildung ihrer Kinder zu investieren. Sie sind respektive sehr engagiert. Im Vergleich zu sozioökonomisch besser gestellten Eltern steht das elterliche Engagement bei sozioökonomisch benachteiligten Eltern jedoch in keinem signifikanten Zusammenhang mit den Lernergebnissen der Schüler:innen.

Zusammenfassend zeigen die Ergebnisse dieser Studie, dass weder die (elterliche) Motivation für Bildungsinvestitionen noch das Ausbleiben von Bildungsinvestitionen ein Problem darstellen. Problematisch ist jedoch die Qualität entsprechender Investitionen. Mit

Blick auf das Apartheid Regime und seine langfristigen Folgen, insbesondere für sozioökonomisch benachteiligte Personen in Südafrika, bleibt die Frage, ob die Verbesserung der Qualität elterlicher Bildungsinvestitionen das Potenzial hat, die Bildungskarriere von Schüler:innen über die Performanz zu lenken; und so der Vererbung sozialer Benachteiligung langfristig entgegenzuwirken. Die Ergebnisse dieser Studie liefern empirisch belegte Hinweise für die Annahme, dass dies der Fall ist, und regen weitere Analysen in diesem Bereich an.

Contents

List of Figures

List of Tables

1. Introduction

Taking into account South Africa's political past and analyzing the country's post-apartheid state, the persistence of supposedly outdated social patterns remains evident. Even though the government aims to advance egalitarian political, judicial and economical agendas, racial segregation and resulting inequalities are deeply rooted in the South African population. It has therefore been posited that a consequence of late Apartheid policy is that South Africa consists of multiple, unequal societies (Haferburg & Osmanovich, 2017).

Aiming at visualizing social patterns of segregation and inequality, photographer Johnnie Miller captured scenes representing the state of socioeconomic inequality in post-apartheid South Africa. The photography included to the right is an example of Miller's project 'Unequal Scenes' (Miller, n.d.). It shows the border between suburbs of Primrose (left) and the informal settlement Makause (right), both located in Johannesburg, South Africa. As shown in the photograph, Primrose and Makause represent contrasting living standards. One result of Apartheid policies of the past is that suburbs like Primrose are mostly inhabited by Whites, and informal settlements like Makause are primarily populated by Black-Africans.

Scientific and artistic depictions of the co-existence of two populations in one country, emphasizing the disadvantage of the Black-African population in South Africa, are also reflected by official statistics. For example, The World Bank issued an assessment report on the state of inequality in South Africa, and summarized that race is still a strong

driver for poverty, and that South Africa's chronically poor population is almost exclusively Black-African and Colored (The World Bank, 2018).

Against this background, the question arises as to which factors are decisive for the persistence of social segregation and inequality. To Max Weber the answer is definite, as he identifies educational attainment to be the main driver of differences in socioeconomic state and life opportunities (R. Becker & Lauterbach, 2016). However, considering theories of rational decision making (Esser, 1999), socioeconomic status is not only determined by educational attainment, but also functions as a predictor thereof (Stocké, 2010). That is, individuals of lower socioeconomic status are less likely to invest in educational attainment due to their cost-benefit calculation. Consequently, upward mobility is unlikely to be achieved, and hence, social inequality remains long-term. Such patterns of educational investment behavior are not only observed among individuals with regard to their own educational career, but especially among parents with regard towards their children's educational attainment (Boudon, 1974; Maaz et al., 2008). This is particularly prejudicial, given that educational non-investments of parents with lower socioeconomic status foster the inheritance, and thus, persistence of social inequality across generations (Esser, 1999). Typically, educational investments are defined as the allocation of monetary resources towards education (Catsiapis, 1987; Stocké, 2010). However, some argue that the factor of time is similarly decisive (G. S. Becker, 1962; Esser, 1999; J. Heckhausen & Heckhausen, 2008). That is, higher educational attainment requires the investment of more time to be successfully completed. Again, along the lines of rational decision making, individuals of lower socioeconomic status are less likely to allocate time towards educational attainment. That is also true for parents' educational investment behavior in the form of parental involvement, and thus, the frequency of involvement activities is considered an educational investment. The relevance of parental involvement for educational attainment has been broadly discussed throughout the literature (Epstein, 1987, 1995; X. Fan & Chen, 2001), and some identify socioeconomic status to be an important predictor for the association of parental involvement and student outcomes in learning (X. Fan & Chen, 2001; Lareau,

1987, 2002; Stevenson & Baker, 1987).

Educational investment behavior has not been explored for the South African context yet, even though the country's highly segregated structure requires an in-depth understanding of the mechanisms and factors that foster the reproduction of social inequalities. Against this backdrop and with the original question of how to mitigate inequalities in mind, this study distinguishes between socioeconomic low- and high-performing groups within the South African population, aiming to explore the degree to which socioeconomic status is decisive to patterns of educational investment behavior. To draw a precise picture thereof, this study utilizes the Rubicon Model of Action Phases (H. Heckhausen & Gollwitzer, 1987; J. Heckhausen & Heckhausen, 2008), and looks at the association between the motivational, volitional, and evaluative phases of actional behavior. To account for the effects of socioeconomic status and parental involvement as a form of educational investment, theories of rational decision making and parental involvement respectively are utilized (Hoover–Dempsey & Sandler, 1997; Hoover–Dempsey et al., 1992; Walker et al., 2005). In doing so, special emphasis is put on separately highlighting phases of human actional behavior, as well as their associations with each other. This approach has not yet been applied to the South African context, and thus, can be considered an original contribution of this study.

The main interest of this study is South Africans of lower socioeconomic status. Typically, socioeconomic status is comprised of financial resources as well as educational and occupational status (Buchmann, 2002). However, as introduced by Bourdieu (1977, 1986) (e.g. Bourdieu & Passeron, 1977) and reinforced by Coleman (1966, 1988) (Coleman, 1988; Coleman et al., 1966), measures reflecting cultural and social capital are important background factors associated with educational performance as well. Thus, many large-scale studies in education account for that by including traditional measures of socioeconomic status as well as additional measures of cultural and social capital; for example, educational resources in the home such as books, access to the internet and availability of a room to one's self (sec. 8.2.3). Hence, this study accounts for a reference group, namely

socioeconomic low-performers, as well as a comparative group, namely socioeconomic high-performers (sec. 8.2.4). Whereas low-performers show lower educational and occupational status and have access to less cultural and social capital, high-performers show the opposite. The qualitative perspective of this study examines a group of socioeconomic low-performers that are adult street vendors, while the quantitative perspective focuses both socioeconomic low-performing and high-performing parents. This terminology is applied throughout this study.

Reflecting on the aforementioned concepts, the following three research questions are addressed in this study: (I) are socioeconomic low-performers in South Africa motivated to invest in education; (II) is the motivation transferred into (parental) educational investments; and (III) are parental educational investments beneficial to student outcomes in learning.

To summarize, this study is motivated by the assumption that education has the power to function as a lever for the enhancement of upward mobility and life opportunities. By exploring the potential of educational investments within and beyond the framework of parental involvement among South African subpopulations, this study aims to add to the understanding of mechanisms affecting educational attainment, and thereby support scientific and governmental efforts to mitigate inequities among socioeconomically disadvantaged groups of South Africa's highly segregated population.

1.1. Structure

To work through the topic exhaustively, this study consists of two parts. The first part focuses the conceptual background, while the second part emphasizes the empirical work of this study. Part one and two consist of the following sections.

Based on Bos (1999), who argues that a country's national context needs to be carefully considered when testing theories across cultural and ethnic groups (Bos, 1999), section

2 provides a broad introduction to South Africa's past and present political, social, and economical state. First, emphasis is put on South Africa's historical past, summarizing the cornerstones of the rise and fall of the Apartheid Regime as well as segregation policies fostering the separation of South Africa's population due to ethnic, cultural and social belonging. Second, the present state of South Africa is described, using the indicators of the Human Development Index (HDI) as a blueprint for structuring this part of the section. Throughout this section, special attention is paid to the differences between socioeconomic low- and high-performers. The overall goal of this section is to derive an evidence-based definition of this study's focus group, namely South African socioeconomic low-performers, as it is decisive for the empirical part of this study.

In section 3, the theoretical framework of this study is elucidated. Based on the grounds of the *Rubicon Model of Action Phases* by Heckhausen & Gollwitzer (1987, 2006), which focused on the sequence of the motivational, volitional, and the evaluative phases of human actional behavior, this study accounts for the separated elements as well as the associations between them. Furthermore, this study utilizes theories of rational decision making introduced by Esser (1999), among others, to contextually enrich the dimension of motivation, and to address the question of why an educational investment is taken or not taken. Lastly, the (revised) *Model of Parental Involvement Process* by Hoover-Dempsey & Sandler (1995, 1997, 2005) and Walker et al. (2005) is introduced, aiming to mirror the theoretical model of general educational investment behavior onto a specific setting, namely, parental involvement.

In section 4, the current state of research is reviewed. Along the lines of empirical contributions to the topic, findings on (parental) educational investment behavior in association with educational attainment and socioeconomic status are systematically summarized and discussed. Finally, the research gaps which this study focuses on, are expounded. Section 5 summarizes this study's research questions as well as corresponding hypotheses with regard to the focus group of this study, namely socioeconomic low-performers in South Africa, and in alignment with the aforementioned theoretical assumptions and empirical state of research.

Lastly, section 6 emphasizes the methodological study design. Although not considered a mixed-methods design, this study utilizes a qualitative and a quantitative perspective. Whereas the qualitative study enables the evaluation of motivation and volition from an individual perspective, the quantitative study allows for analysis of parental involvement activities as well as their association with student outcomes in learning. Hence, this study makes use of two separate methodological approaches, both aiming at enhancing the understanding of educational investment behavior among socioeconomic low-performers in South Africa.

Following the conceptual background, the second part emphasizes the empirical work of this study. First, section 7 emphasizes the qualitative perspective. Qualitative data was collected in Cape Town/South Africa in 2013, in a study on educational needs of street vendors. Hence, the sample includes adults of lower socioeconomic status, namely socioeconomic low-performers. On that note, the qualitative study takes a broader perspective on the topic of educational investment behavior, aiming at evaluating whether motivation for educational investments is evident among the focus group, and whether it is put into action. The qualitative study serves as an opening for the sequential quantitative study. Subsequent to the qualitative perspective, section 8 focuses the quantitative perspective of this study. Whereas the qualitative perspective emphasized educational investment behavior among socioeconomic low-performing adults, the focus of the quantitative perspective is to evaluate educational investment behavior among socioeconomic low-performing parents, and whether their involvement activities are positively associated with student outcomes in learning. Data used for this part of the study come from IEA's prePIRLS assessment of the *Progress in Reading and Literacy Study* (PIRLS) 2011,[1] providing a

[1] The International Association for the Evaluation of Educational Achievement (IEA) is an international non-profit organization constituted of national research institutions, government research agencies, scholars and analysts aiming at contributing to the evaluation and improvement of education and educational policies worldwide. The IEA professionalized in planning, conducting, and leading international large-scale assessments measuring student outcomes in reading, mathematics, and science among others since 1958. Further information are available at: www.iea.nl

nationally representative sample of fourth-grade students and their parents.
Lastly, results as well as limitations and advantages of the qualitative and quantitative study are discussed in section 9, followed by section 10, which provides a brief summary of the study as well as its implications and points of connection for further research.

1.2. Terminology

The following section sheds light on terminologies used throughout this study. Generally, labels for groups identified for this study are adapted from other publications and official reports, aimed at utilizing the appropriate terms. This is also along the lines of current movements towards social responsibility, inclusive societies, as well as *Othering and Belonging* (e.g., Powell, 2012).

"Racial terminology in South Africa is a complicated matter" (Seekings & Nattrass, 2005, p. ix). This statement proves to be true when considering official and scientific work on the South African population, and the diverging application of terms representing its subpopulations. On that note, an appropriate appellation is necessary to emphasize South Africa's political and historical past, and present state.
Throughout official statistics published by the Republic of South Africa, department of Statistics South Africa (Stats SA), the terms *Black-African* or *African*, *Colored*, *Indian or Asian* and *White* are applied (see Republic of South Africa, 2012a). Black-Africans or Africans are represented by individuals originating from native African groups such as Zulu, Xhosa, Ndebele and Swazi people as well as the Tswana people among others. Coloreds are represented by individuals descending from mixed origins, and Whites are represented by individuals descending from European origins, mainly speaking English or Afrikaans (Republic of South Africa, 2012b). An alternative key for the distinction of ethnic groups in South Africa uses *African* for those classified as *Native*, *Black* or *Bantu* by the Apartheid State; *White* is used for those originating from Europe; *Indian* is used

for those originating from India;[2] and *Colored* is used for those not fitting any of the other categories. Respective distinctions are often limited to White and Black, the latter referring to "African, Colored, and Indian people collectively" (Seekings & Nattrass, 2005, p. xi). Yet another approach is to distinguish even more broadly between ethnic groups and refer only to *Whites* and *Non-Whites* (Junge, 2017).

Based upon the application of terms briefly reflected above, this study uses two approaches. First, a distinction is made between *Whites*, which are individuals primarily originating from Europe during colonialism, and *Non-Whites*, which are Black-Africans, Coloreds, Indians and Asians (see Republic of South Africa, 2012b). Given that the latter collectively suffered from the policies of the Apartheid Regime, they are considered a contrasting group to Whites. However, this broad distinction of two groups is only applied when a more distinct breakdown does not reveal further information on the specific topic of inquiry. However, to make a statement applicable to a particular group only, this study uses the terms *White*, *Black-African*, and *Colored* following the terminology applied in official governmental reports (see Republic of South Africa, 2012a). The group of South Africans referred to as as Colored summarizes Indians, Asians, as well as all others not fitting the categories of White or Black-African people.

In addition to ethnic groups, distinctions between different areas of living are elementary to the South African population as well. Along the lines of official reports, this study distinguishes between urban and suburban areas, remote and remote rural areas, and townships (e.g., Howie et al., 2012, 2018).

Townships are a particularity of the South African spatial structure and emerged during the Apartheid Regime to home the Non-White population. They are still located in the close vicinity of larger cities and not infrequently connect directly to the former White residential areas. However, as shown in the photographic work of J. Miller, townships typically show vestigial infrastructure (Jürgens et al., 2013). Another expression

[2] They were also referred to as *Asiatic* by the Apartheid State (Seekings & Nattrass, 2005).

of segregation policies during the Apartheid Regime yielded the creation of *Homelands*. Originally known as *Bantustan Land* or *Black States*, Homelands were designated areas across the country where the Non-White population was forcibly allocated. Considered citizens of the Homelands, the Non-White population also lost their South African citizenship (Britannica, n.d. Platzky & Walker, 1985). Up to today, former areas of Homelands show high levels of poverty (The World Bank, 2018). Homelands are commonly understood as a euphemistic description of the underlying concept of suppression and racial discrimination. Therefore, the term is not commonly used in the literature, and thus, is not applied in this study unless needed for the appropriate description of political and historical events of the past.

2. South Africa in Profile

Given the country's history and its inevitable effect on the societal structure of today, it is essential to focus on both South Africa's past and present. To that end, the first part of this section outlines the corner stones of the Apartheid Regime. The second part focuses South Africa's present. First, an overview of the demographic structure is provided. Second, the dimensions of the Human Development Index (HDI) are utilized as a blueprint for a comprehensive presentation of the country's current status.

2.1. South Africa's Past: A Retrospective on Apartheid

2.1.1. Definition of Apartheid

In the 1990s, Archbishop Desmond Tutu called South Africa the *Rainbow Nation*, stressing the cultural and ethnic diversity of the country's population. However, political agendas of the past aimed at disabling multiculturalism and carried out authoritative policies of racial segregation widely known as the Apartheid Regime (Republic of South Africa, 2012c). Originating from the Afrikaans meaning of *apartness*, Apartheid refers to separateness or separation. However, in the South African context, Apartheid represented a policy of racial segregation (Bunting, 1972). More precisely, it was defined as the attempt of a governmental manifestation of racism, and hence, was an instrument for manipulating and restructuring society. During a time of modernization and economic growth tailored to serve South Africa's White population, Apartheid represented exclusion, suppression and exploitation of the Non-White, mostly Black-African population (Marx, 2017). The Apartheid Regime is therefore understood as a bilateral system "securing the basis of white privilege and superiority" while suppressing the Non-White population (Reddy, 2004, p. 9).

2.1.2. Political Events heading the Rise and Fall of Apartheid

The prerequisites of Apartheid were created in the aftermath of colonialism in the early 20th century, starting with the unification of settlements owned by the British Empire

and the Boer Republic in 1910. Also known as the *Act of Unification*, the merger led to a conglomerate of different political and cultural facets, mostly headed by former settlers. With the ideals of colonialism still internalized, the newly created community was influenced by an artificially created hierarchy in which three aspects, namely color of skin, cultural belonging and social status, were highly correlated (Haferburg & Osmanovich, 2017). White, wealthy settlers originating from Western communities kept reign over the Non-White, mostly Black-African population, continuously practicing their suppression, slavery and exploitation, and thereby establishing an "ideology of White supremacy and Black inferiority" (La Guma, 1972, p. 12). Essentially being a *White Union* (Republic of South Africa, 2012c), the newly founded South African society deprived its Non-White population of the chance to recover from the repercussions of colonialism. Instead, the Non-White population suffered from severe subjugation, being "landless, voteless and compelled to provide cheap labour to produce the fabulous wealth" of the White population (La Guma, 1972, pp. 13–14). With the government paving the way towards Apartheid, the African National Congress (ANC) was founded in 1912 as a reaction to the continual discrimination of the Black-African population (Republic of South Africa, 2012c).

The development reached its climax in 1948 when the pro-Afrikaner National Party (NP) came to power and put the subliminally practiced ideological conviction of racial segregation into a political frame, known as the Apartheid Regime (Haferburg & Osmanovich, 2017). Thus, it was concluded that the Apartheid policies led South Africa into "a colonial situation (...), more precisely an internal colonialism" (Marx, 2017, p. 45). Reacting to the political movements of the time, the ANC stressed its *Programme of Action*, advocating for "a rejection of white domination and a call for action in the form of protests, strikes and demonstrations" (Republic of South Africa, 2012c, p. 4).

In the early 1950s, the ANC Youth League (ANCYL),[3] which emerged from the ANC, carried the *Defiance Against Unjust Law Campaign*, protesting against racial segregation in South Africa. At the same time, the government adopted the *Bantu Education Act*

[3] The ANCYL fostered political figures such as Oliver Tambo, Walter Sisulu and Nelson Mandela (Republic of South Africa, 2012c)

fostering the nationalization of the educational system.

In the 1960s, oppositional parties such as the ANC were banned and their political initiators were sentenced to life imprisonment. After winning a *Whites-Only Referendum*, the NP declared the Union of South Africa a republic, actively practicing the division of the Rainbow Nation into artificial ethnic nations (Republic of South Africa, 2012c). The government was not hesitant to use physical violence, that is systematic torture of the Non-White population, to ensure the continued existence and success of Apartheid's ideals (Marx, 2000). While the global trend increasingly led towards the abolition of corporeal punishment, South Africa was among few countries around the world where the opposite was the case. By the end of the 1980s, the South African government still sentenced 40,000 people to fustigation per year (Marx, 2017). Approximately 86% of the Black-African and 72% of the Indian population were physically tortured during Apartheid. On the contrary, only 23% of the White population reported physical abuse during that time (Marx, 2000).

It was only in the mid 1970s that the Apartheid Regime reached its tipping point. Due to industrial and economical developments, the ideals of Apartheid, namely racial segregation and suppression of the Non-White population, were no longer compatible with the increasing demand for a well-trained, well-qualified and long-lasting workforce. Thus, instead of a rigorous distinction of groups according to skin color, cultural belonging and social status, privileges were now given to those who had work. Due to recent events of the past, the majority of Black-Africans still struggled to obtain gainful employment (Marx, 2017). Forced by the need of a functioning workforce, the NP agreed on a set of reforms in the 1980s, discontinuing certain acts of discrimination. Also, the international community expressed their support for oppositional parties such as the ANC, and pressured the South African government for change of course (Republic of South Africa, 2012c).

The path towards the end of Apartheid was paved in 1990 when President F.W. de Klerk announced the legalization of liberation movements initiated by the ANC, among others, and released political figures such as Nelson Mandela from life imprisonment (Marx, 2017;

Republic of South Africa, 2012c). In 1994, South Africa voted the ANC to be their leading party, and Nelson Mandela became the first democratically elected president (Republic of South Africa, 2012c).

2.1.3. Segregation Policies during Apartheid

Capturing the motives of South Africa's White population for pursuing Apartheid policies, the journalist Antjie Krog (1998) once highlighted the following statement from a policeman and Apartheid attendant:

> "We believed black people were not human; they were a threat, they were going to kill us all, and then waste away the country until it was nothing but another African disaster area. (...) While some men were out killing black people, many whites were busy dreaming of a life without black people: separate laws, separate amenities, separate churches, separate homes, separate towns, separate countries." (Krog, 2000, p. 93)

Complementing the White perspective, a victim of Apartheid emphasized the Non-White experience. During an interview he said that he cannot understand how one could be forced to leave behind their homes, work and friends. "He can only understand it if he says to himself that he is not a man. But then, what is he?" (Platzky & Walker, 1985, p. 6).

Illustrating human perceptions during Apartheid, both statements noticeably emphasize the aspects of separation and *apartness* due to race. Being the core tenet of Apartheid, racial segregation was realized by different measures, above all residential and educational separation leading to severe disparities in opportunities of social mobility. Both aspects are highlighted throughout the literature and are highlighted in the following section as well.

Fostering the elevation of the *White Man's Places*, the NP aimed at counteracting the movement of the African population, which was increasingly settling in city centers (Marx,

2017). It is estimated that between 1960 and 1983, approximately 3.5 million Black-African and Colored people were affected by relocation policies and forced removals initiated by the South African government (Marx, 2017; Platzky & Walker, 1985).[4]

To ensure racial segregation, the NP established different measures, one of which were pass laws. Already in place since the early 1800, pass laws obliged every African to carry a passport that held personal information as well as the current place of employment. Those not working in the city centers were required to leave after a maximum of 72 hours. The Non-White population was highly affected since the majority worked in farming and mining, mostly located in the rural areas of the country. Moreover, the pass laws enabled the government to actively manipulate the labor market, ensuring that the Non-White population was not employed in the city centers. It was only in 1986 that pass laws were finally abolished (Marx, 2017).

Further, the acquisition of land by the native African population was largely prohibited. Already initiated in 1913 with the *Natives Land Act*, the government granted an exception of 13% in 1936, aiming to allocate space for the removals of the Non-White population from the city centers (Marx, 2017). The so-called Homelands, which were designated reserves in rural areas (sec. 1.2), were planned to be developed into self-governed units. However, it was evident that respective areas would be economically unviable due to missing resources and, therefore, would always depend on the South African government. By forcing the removal of millions of Non-Whites from White areas and relocating them to the Homelands, these areas were commonly overpopulated, causing impoverishment and the expansion of diseases such as HIV and Tuberculosis. By establishing the Homelands, the government further managed to ensure reservoirs of cheap labor and thereby exacerbate the suppression of the Non-White population (Marx, 2017).

In addition to the Homelands, the NP fostered the creation of townships. Still home to millions of Africans in today's South Africa, townships are characterized as dumpsters of

[4] The number of people being affected by relocation policies during Apartheid can only be estimated since no reliable records exist of those who were victims of forced removals in urban areas, and those who were affected by moves within the reserves (Platzky & Walker, 1985).

surplus people, namely people that were not considered to be of use to ensure the functionality of South Africa's economic growth. Equal to the situation in the Homelands, townships were heavily overpopulated and suffered from missing infrastructure such as electricity, running water and wastewater disposal (Marx, 2017).
While the cities became centers of development, modernization and economic growth, Homelands and townships were deliberately underdeveloped, excluding their citizens from modern infrastructure, medical support and transportation systems, including streets and trains (Marx, 2017). Built on the bilateral ideology of the Apartheid Regime, this development pointedly exemplifies its contrary development, that is modernization and growth among Whites versus suppression and racial segregation among Non-Whites.

Besides acts of relocation, the Apartheid policies of racial segregation also affected the educational system of South Africa. As well as areas of living, the educational landscape was subjected to the *apartness* of White and Non-White schools.[5]
Rooted in times of colonialism, disparities with regard to educational opportunities between Whites and Non-Whites persisted in the South African population for a long time. Mainly carried out by private and church authorities, schools generally focused on teaching English as well as Western values and ways of life, though, very few Non-Whites had access to education. With the NP coming to power in the late 1940s and introducing the *Bantu Education Act* in 1953, the educational system was subjected to nationalization, and the South African government aimed at ensuring control thereof (Junge, 2017). With the Act, educational matters were handed over from provincial control to the Ministry of Bantu Affairs, headed by South Africa's future Prime minister Hendrik Verwoerd (H. Bernstein, 1972).
Placing Bantu education in the light of Apartheid's premises, Verwoerd emphasized the presumed superiority of the Whites, declaring the following:

[5] No in-depth distinction of primary, secondary, and higher education is made here, given the irrelevance of such for the purpose of the conceptualization of the South African Non-White population during Apartheid.

> "Native education should be controlled in such a way that it should be in accord with the policy of the state. (...) [Further,] it is of no avail for him [the White man] to receive training which has as its aim absorption in the European community. He must not be subject to a school system which drew him away from his own community and misled him by showing him the green pastures of European society in which he is not allowed to graze." (cited after H. Bernstein, 1972, pp. 45–46)

He continued by addressing the main purpose of Bantu education, namely to teach skills and knowledge that are primarily beneficial to the community and prepare the individual to serve the needs of South Africa's economy. Resulting from the separation of the role of Non-Whites in South Africa's community as well as the main purpose of their education, Verwoerd summarized that there is no point for them to learn apart from what is necessary for the survival of the state. "What is the use of teaching the Bantu child mathematics when it cannot use it in practice? (...) Why should girls bother with (...) art, or drama, or literature. They could have babies without that sort of knowledge" (cited after H. Bernstein, 1972, p. 46).

The Bantu Education Act therefore served two purposes in addition to ensuring governmental control over the educational system. These purposes are, first, the protection of the Apartheid Regime; and second, the protection of the wealth of the White population (e.g., H. Bernstein, 1972; Junge, 2017). Therefore, the NP implemented a curriculum that focused on teaching Western values as well as stereotypical social roles of Whites and Non-Whites. For example, Non-Whites were taught that their *otherness* was natural. The creation of mainly non-resistant, docile individuals was sought (Reddy, 2004). To achieve such goals, different measures were taken.

For example, access to education was limited for Non-Whites. Education was only partly compulsory for Coloreds and not compulsory for Black-Africans. Drop-out rates for the latter were exceptionally high. In 1968, the majority of Black-African students (70%) attended the first four years of schooling, but only one out of three would attend more than five years (H. Bernstein, 1972). Drop-out rates were due to many reasons, including

the concept of academic ceilings, which was only practiced in Non-White schools. Those identified respective thresholds that a student had to successfully reach in order to be permitted to continue an educational career. Students not reaching the academic ceiling were considered to be not "mature enough or gifted enough to derive further benefit from academic studies" (cited after H. Bernstein, 1972, p. 49). They were excluded from further education accordingly. Hence, drop-out rates in Non-White schools were essentially artificial drop-outs or, more precisely, deliberate eliminations (H. Bernstein, 1972).

Another measure was to distribute educational resources unequally among White and Non-White students. In 1965, 77% of the annual national governmental allocations in education was spent on the education of Whites. In comparison, only 8.9% was dedicated to the education of Black-Africans. Considering the population sizes, respective spending was equal to ZAR74 ($4) per capita on White children, and less than ZAR3 (<$1) per capita on Black-African children.[6] Education for Whites was free and included the provision of school facilities, staff, and learning material such as textbooks and writing utensils. In comparison, Black-Africans were required to pay school and examinations fees and provide for learning materials such as textbooks, school transportation, and uniforms themselves. Additionally, Black-African schools and parents accumulated monetary resources for additional teacher salaries ensuring sufficient lessons for all children in Black-African schools. In 1967, 20% of all teachers were paid by private funding (H. Bernstein, 1972).

Furthermore, quality of education was intentionally suppressed. Under the guise of the Bantu Education Act of 1953, double-session schools had been established. The concept required teachers in Black-African schools to teach two to three sets of children in one day, each session lasting for three hours. Thereby, the government aimed at educating the

[6] The South African currency is the South African rand, or simply rand (ZAR). In line with APA style, the official code is used to indicate the currency of a monetary value. Throughout the study, both U.S. dollars (USD; $) and South African rand (ZAR) are used. In the case of values being reported in ZAR, they are converted to USD as of August 2022, and rounded to the full number. Though, conversion rates have changed over time implying that ZAR might have equaled +/- USD at the point in time that it was originally documented. Hence, respective conversion rates merely serve as a point of reference.

large number of Non-White children without the need to enlarge the education fund. The concept was unknown to White schools. In 1969, double-sessions were practiced in 51% of primary Non-White schools (H. Bernstein, 1972). Furthermore, even though conditions of the teaching profession were tough, Bantu teachers were not adequately paid. To justify that, Verwoerd claimed that Bantu teachers were required to "serve the Bantu community. (...) [The] salary must be fixed accordingly. (...) The salaries which European teachers enjoy are in no way a fit or permissible criterion for the salaries of Bantu teachers" (cited after H. Bernstein, 1972, p. 54). That being said, Bantu teachers became instruments of Apartheid policies, operating between limited access to resources, insufficient salaries, and a dictated curriculum ensuring to keep quality of education at a lower level.
Lastly, language policies were implemented, ensuring the goals of Apartheid's Bantu education and adding to a lack of quality thereof (Junge, 2017). Prior to the Bantu Education Act, teaching was conducted in respective mother tongues only until grade 4 or 5. However, after 1953, Non-White students received instructions in their native language throughout all grades. However, since English and Afrikaans were accounted for as the only two official languages, all governmental, economical, and environmental communication was carried out in one of these two languages (H. Bernstein, 1972). Hence, by purposely educating Non-Whites primarily in their native language while simultaneously practicing all official and governmental affairs in either English or Afrikaans, the NP deliberately fostered the manifestation of segregation and inequality (H. Bernstein, 1972; Junge, 2017).

2.2. South Africa's Present: The Aftermath of Apartheid

The post-colonial ideals as well as the Apartheid Regime led to severe disparities in life-expectancy, education, and standard of living, causing long-term inequality of social mobility among South Africa's population (Marx, 2017). Against this background, the following section addresses the current state of South Africa, and the key dimensions of

the *Human Development Index* (HDI) are utilized as a blueprint to structure this section.[7] In the latest round of assessment in 2019, the UNDP assigned South Africa an HDI of 0.709.[8] At position 114 out of 189, South Africa is placed at the lower end of countries assigned to the category of high human development (UNDP, 2020). Since 1990, South Africa's HDI increased by 0.42 points,[9] or approximately 13% (UNDP, n.d.). Between 2011 and 2013, that is the time frame of this study, the South African HDI increased from 0.665 to 0.685.

In addition to the HDI, this study utilizes South African census data of 2011 to report on the present state of South Africa.[10] Respective data is sourced from the Department of Statistics South Africa (Stats SA).

2.2.1. Demography

As of 2011, South Africa is populated by approximately 51.8 million people. About 51% of the population is female, and 49% is male. South Africa is a fairly young country, considering that in 2011, about 59% of the population is aged below 30 (Republic of South Africa, 2012a).

Approximately 79% of the South African population is classified as Black-African, and

[7] First introduced in 1990 by the United Nations Development Programme (UNDP), the HDI is an annual summary measure of three key dimensions aiming at mirroring human development. These key dimensions are health and life expectancy, education, and standard of living, all of which are accompanied by respective sub-dimensions. Health and life expectancy are based on the life expectancy at birth; education is based on the mean years of schooling for adults aged 25 or older and the expected years of schooling for children at school entry age; and standard of living is measured by gross national income per capita (GNI) (UNDP, n.d.). Furthermore, key dimensions of the HDI are linked to the United Nation's *Sustainable Development Goals* (SDGs). Agreed on by countries around the world in 2015, the United Nations developed the *2030 Agenda for Sustainable Development*, aiming at motivating efforts to end poverty, inequality, and climate change around the world until 2030. In line with respective efforts, 17 *Sustainable Development Goals* (SDGs) have been comprised, ensuring opportunities for all. The SDGs are a follow-up of the *Millennium Development Goals* (MDGs) (United Nations, n.d.). The dimension of health and life expectancy is linked to SDG 3 (*Good Health & Well-Being*); education relates to SDG 4 (*Quality Education*); and lastly, standard of living relates to SDG 8 (*Decent Work & Economic Growth*).

[8] In 2019, the HDI ranged between 0.957 in Norway and 0.394 in Niger.

[9] The value is computed as the annual compound growth rate of the HDI (UNDP, 2020, p. 350).

[10] The latest cycle of the population census has been administered in early 2022, though respective results are not published yet.

hence, which constitutes the majority of the South African population. The second largest groups are Coloreds and Whites, each constituting 8.9% of the population. Indians and Asians account for 2.5%, and other ethnicities account for 0.5% of the South African population (Republic of South Africa, 2012a).

South Africa recognizes 11 official languages, namely English and Afrikaans as well as several native African languages.[11] The most frequently spoken languages in South Africa are IsiZulu (22.7%) and IsiXhosa (16%), followed by Afrikaans (13.5%) and English (9.6%). Looking only at the province Western Cape, where Cape Town is located, IsiXhosa is frequently spoken by 24.7% of the population, which makes it the second most frequently spoken language after Afrikaans (49.7%) (Republic of South Africa, 2012a). Although English is only spoken most frequently by approximately 10% of South Africans, it is the country's most used and understood language (Republic of South Africa, 2012b). However, looking at population groups, Black-Africans most frequently speak IsiZulu (28.5%) and IsiXhosa (20.1%), and only 2.9% declare English to be their first language (Republic of South Africa, 2012a).

2.2.2. Health

According to the UNDP, the life expectancy at birth in South Africa was 58.9 years in 2011, and 64.1 years in 2019 (UNDP, n.d.). Although life expectancy at birth increased by approximately 5 years between 2011 and 2019, South Africa still deals with several health issues at present. Above all are Human Immunodeficiency Virus (HIV) and Tuberculosis (TB) exacerbate by the segregation policies of Apartheid that perpetuated bad living conditions and severe poverty of the Non-White population.

According to UNAIDS, South Africa has the world's highest HIV rate. Latest data of 2021 showed that over seven million South African adults and children live with HIV. The prevalence rate for people aged 15 - 49 was estimated at 18.3% (UNAIDS, 2021).

[11] The 11 official languages of South Africa are English, Afrikaans, isiNdebele, isiXhosa, isiZulu, Sepedi (Sesotho sa Leboa), Sesotho, Setswana, siSwati, Tshivenda, and Xitsonga (Republic of South Africa, 2012b).

The government took several counter measures trying to prevent new infections. Among others, South Africa has the largest antiretroviral therapy (ART) program in the world (Republic of South Africa, 2012d). 74% of South African adults and children infected with HIV received ART in 2021 (UNAIDS, 2021). A particularity of HIV treatment in South Africa is the combination with TB. Although measures to controlling TB are improving, the cure rate has not reached the international standard of 85% recommended by the World Health Organization (WHO) by 2012 (Republic of South Africa, 2012d). In 2017, TB was the leading cause of death among the age groups 15 - 44 and 45 - 64, followed by HIV which was ranked second and third respectively in those age groups (Republic of South Africa, 2017).

Another major concern of the South African health sector is hunger. The South African National Health and Nutrition Examination Survey found that in 2013, 54.3% of South African households suffered from hunger; that is, they were either at risk of hunger (28.3%) or faced food insecurity (26%). Both factors primarily affected the Non-White population, which showed the highest scores in both categories (Shisana et al., 2013). On the contrary, the General Household Survey carried out by Stats SA found that vulnerability to hunger only affected 11.6% of households in 2011. The numbers remained somewhat steady after that point, remaining at 11.6% in 2021 (Republic of South Africa, 2021). In summary, numbers on the current status of food security are limited and deviate severely from each other. However, it was found that in 2016, malnutrition was among the ten leading causes of death for infants in South Africa (3.2%), as well as children and teenagers aged 1 - 14 years (3.7%) (Republic of South Africa, 2017).

A major step towards equity among all ethnic groups of the South African population is the goal of a National Health Insurance (NHI). Released in 2011, the proposal aimed for universal coverage of care across all areas. Although not built to include all health services, the NHI aims at providing "essential healthcare to all citizens of South Africa (...) regardless of their employment status and ability to make a direct monetary contribution to the NHI Fund" (Republic of South Africa, 2012d, p. 255). In 2017, the NHI became official. However, in 2015 the World Bank reported that the high-income population lived

significantly closer to health facilities such as hospitals than the low-income population, and thus, access to medical care remains troublesome (The World Bank, 2018).

2.2.3. Standard of Living

In order to estimate the standard of living and indicate a country's economic strength, the UNDP draws on the *Gross National Income* per capita (GNI). In 2012, South Africa showed a GNI per capita of $12,404, which increased to $12,129 in 2019 (UNDP, n.d.). For comparison, the highest ranking country by this estimate, Norway, showed a GNI per capita of $66,494 in 2019. According to the World Bank, Norway is considered a country of very high human development, whereas South Africa is among the countries considered to have high human development (UNDP, 2020). Thus, The World Bank positions South Africa among the 'upper middle-income' countries,[12] although poverty rates remain rather high (The World Bank, 2018). This can be partially explained by the fact that the South African population consists of few high-income earners, and many low-income earners due to segregation policies of the past. While low-income earners are large in number, they generate relatively little of the South African annual household. More precisely, the top 10% of earners held about 70% of the net wealth in 2015, while the bottom 60% held only 7% (The World Bank, 2018).

According to the World Bank, the international poverty line is set to at or below $1,90 per day. The poverty line of countries classified as lower middle-income countries is set to $3,20 per day, and to $5,50 per day for countries classified as upper middle-income countries.[13] As of 2014, about 19% of South Africans lived on $1,90 per day, 37% lived on $3,20 per day, and 57% lived on $5,50 per day (The World Bank, n.d.-c).

Referring to results summarized in the *Poverty and Inequality Assessment Report* published by the World Bank and Stats SA in 2018, "nearly half of population of South Africa

[12] Using the World Bank Atlas method. Further information are available at (The World Bank, n.d.-a).

[13] Originally fixed to a standard of $1 per day in 1985, the threshold of the international poverty line has been risen to $1,90 per day in 2015. The poverty headcount ratios per day are the percentages of the population living on less than the predefined threshold per day at 2011 international prices (The World Bank, n.d.-b).

is considered chronically poor", earning less than ZAR992 ($58) per month in 2015 (The World Bank, 2018, p. xviii).

Race is still one of the strongest indicators related to poverty, and the group of chronically poor in South Africa consists almost exclusively of Black-Africans and Colored citizens (The World Bank, 2018). In general, rural areas have higher poverty rates than urban areas. In 2015, almost 60% of the poor lived in rural areas and suffered from overcrowded living conditions.[14] Although access to electricity, sufficient water supply, as well as sanitation facilities improved continuously since 1994 and became somewhat universal, official statistics show that in 2011, 99.3% of the White population had access to piped water, compared to 89.1% of the Black-African population (Republic of South Africa, 2012a; The World Bank, 2018).

Work Situations

Associated with the remaining levels of poverty and inequality is the persistence of high unemployment rates. Particularly affected is the Non-White population as well as women, who are likely to struggle finding work and receiving sufficient payment. Furthermore, unemployment rates among the younger population aged 20 - 29 were estimated at about 40% for the 2005 - 2015 period (The World Bank, 2018). This is particularly concerning considering the fact that the majority of South Africa's population is aged below 30 (sec. 2.2.1). Hence, a large share of the South African population participating in the labor force are considered to be working poor. Furthermore, a major issue is the poor state of the education system, causing lack of skills among untrained workers (The World Bank, 2018).

The informal sector is therefore important to the South African economy. In search of a comprehensive definition, Heintz & Posel (2008) draw on different concepts proposed by the *17th International Conference of Labor Statistics* (ICLS), and propose that the informal sector must refer to "(1) self-employment in informal enterprises and (2) em-

[14] A standard of a two persons-per-bedroom is considered in order to rate whether living space was overcrowded (The World Bank, 2018).

ployment in informal jobs" (Heintz & Posel, 2008, p. 27). Furthermore, the definition must also acknowledge informal work outside the informal sector, for instance, unregistered occupations in private households (e.g., domestic work) or individuals who are not employed in formal businesses, have no employment contract, and do not have the right to paid absences or pensions (e.g., street vending). Finally, the authors refer to "formal as protected (or regulated), and informal (...) as unprotected (unregulated) employment" (Heintz & Posel, 2008, p. 32). In line with the proposed definition, Stats SA understands informal work as such that is not registered for income and value added tax, secured by a contract of employment, or contributing to medical insurance and pension funds (Republic of South Africa, 2018).

Based on the *South African Quarterly Labour Force Survey* (QLFS) 2018, the WIEGO[15] reports that about 30% of total employment (including agricultural and non-agricultural labor) in South Africa is located in the informal sector.[16] Furthermore, 48% of employment in rural areas is considered informal work, compared to urban or metropolitan areas where informal labor represents about 24% of total employment.[17] The QLFS of 2010 showed that women (29%) work more frequently in the informal sector than men (23%). Furthermore, Black-Africans are more likely to be involved in informal labor, compared to Whites and Indians (The World Bank, 2018). In general, there are three major fields of work within the informal sector, namely domestic work and private household services, street vending, and market trading. Occasionally, waste-picking is considered as an additional category (Budlender, 2011; Rogan, 2019). Official statistics of the past decade show that private household services remain the largest share of the informal sector, followed by street vending. The latter was estimated to make about 15% of the informal sector in 2007 (Wills, 2009). In total, employees as well as self-employed workers in the

[15] Women in Informal Employment: Globalizing and Organizing (WIEGO) is a global network aimed at improving the working conditions of the working poor in the informal economy with a special focus on women. Further information are available at: www.wiego.org

[16] However, based on respective statistics, the WIEGO informs that only 2% of the total informal labor in South Africa falls onto the agricultural sector (Rogan, 2019).

[17] According to the WIEGO, respective numbers are estimated based of the total employment within the eight major metropolitan areas of South Africa (Rogan, 2019).

informal sector earned about ZAR10 (<$1) on an hourly rate.[18] On the contrary, work in the formal sector yielded around ZAR22 ($1) per hour for both, wage employees and self-employed workers, equaling a difference of more than 100% in favor of workers in the formal sector (Wills, 2009).

Respective income disparities in the South African labor market are not solely due to obvious characteristics such as education, skills and knowledge or work experiences, but also due to entrance and mobility barriers (Heintz & Posel, 2008). A vast majority of studies also showed that private capital for starting a business as well as language proficiency in English or Afrikaans potentially enhances employment opportunities (Heintz & Posel, 2008).

Governmental Efforts

To counteract the precarious situation of the national labor force as a predictor of poverty and social inequality, affirmative and fiscal governmental programs were established. Among those are the National Health Insurance (sec. 2.2.2), and the social wage program. The latter refers to governmental investments in a broad range of societal areas, for example education and infrastructure (The World Bank, 2018). Another prominent example is the *Broad-Based Black Economic Empowerment* Act of 2003 (BBBEE). The BBBEE supposedly fosters "economic empowerment of all black people including women, workers, youth, people with disabilities and people living in rural areas through diverse but integrated socio-economic strategies" (Republic of South Africa, 2004, p. 4). Thus, the BBBEE Act attempts to pave the way for increasing participation of the Non-White population in managerial positions and ownership, and thus, aims at counteracting remaining levels of socioeconomic inequality, fostered through poverty and income deprivation (Patel & Graham, 2012). For example, the BBBEE would assign credits to companies who ensure a certain share of Non-White ownership or make use of supply chains involving companies that meet the policy's requirements (The World Bank, 2018).

[18] Wage employees earned about ZAR6 (<$1)per hour which is slightly less than self-employed workers, who earned ZAR9 (<$1) per hour working in the informal sector in 2007 (Wills, 2009).

Although research on the effectiveness of respective governmental efforts is limited, for example due to lack of an effective monitoring and reporting system (Shai et al., 2019), some findings point towards beneficial effects on the transformation of societal patterns and the rise of a Black-African middle class (e.g., Horwitz, 2011; Patel & Graham, 2012). For example, research utilizing data collected by Stats SA at several points in time (1996, 2001, 2007, 2011) found that the share of the Black-African population being exposed to deprivation has significantly decreased over time (Burger et al., 2015).[19]
However, findings also confirm the persistence of "racial poverty dominance" majorly affecting the Black-African population (Burger et al., 2015, pp. 12–13). That is, Black-Africans are continuously exposed to higher levels of deprivation than any other societal group of the South African population. Hence, South Africa is still considered one of the most unequal countries in the world (Burger et al., 2015; Spaull, 2019).

2.2.4. Education

The South African School System

Since the Apartheid Regime came to an end in the mid 1990s, the educational system of South Africa was subjected to major changes. In 1996, Bantu Education practiced during Apartheid was replaced by the *South African School Act* (SASA). Based on the SASA, the Department of Basic Education (DBE) made education compulsory between the ages 7 - 15, or until grade 9 respectively (Howie et al., 2012).[20] To that end, the DBE involved different entities essential to the success of compulsory education. Besides mobilizing parents who are responsible for sending their children to school, this entailed the accessibility and affordability of education ensured by the South African government as well.

[19] The index of deprivation consists of information on overcrowding, dwelling type, the main source of energy for cooking, water access, telephone/mobile phone access, refuse removal, sanitation, employment, and educational level (Burger et al., 2015).

[20] In 2009, the DBE was separated from the *Department of Higher Education and Training* (DHET) (Howie et al., 2012). Whereas the DBE looks after primary and secondary schooling, the DHET is responsible for post-secondary and tertiary education as well as vocational training.

The present South African education system consists of three levels, that is *General Education and Training* (GET), *Further Education and Training* (FET), and *Higher Education and Training* (HET) (Houserman, 2016; Junge, 2017; Republic of South Africa, 2012e). GET covers grades 1 - 9, whereas grades 1 - 7 are classified as primary schooling, and grades 8 - 9 are considered secondary schooling. GET is further subdivided into three levels. Level I *Foundation* covers grade 1 - 3; level II *Intermediate* covers grade 4 - 6; and level III *Senior* accounts for grade 7 - 9. Level I is tailored to educate students aged 7 - 9 years, level II addresses students aged 10 - 12, and level III focuses on students aged 13 - 15 years. Furthermore, FET covers post-secondary education, grades 10 - 12, and results in the acquirement of a university entrance qualification. The *National Senior Certificate* (NSC), also known as *Matriculation* or *Matric*, entitles students to continue with Higher Education (HE). Finally, HE comprises all tertiary qualifications, namely university degrees (Houserman, 2016). Additionally, the DBE implemented compulsory pre-primary education, aiming at including all children to a reception year, grade R, prior to entering general education and training. Pre-primary education has been in place since 2012 (Howie et al., 2012).

The South African school system distinguishes between public and independent schools. The latter are also considered private schools. According to the SASA of 1996, public schools are established, funded and maintained by the respective provincial government whereas independent schools are subjected to private administration and funding. However, independent schools are required to meet the regulations established by the South African government in order to qualify for registration (Republic of South Africa, 1996). In total, 3.8% of South African schools are independent schools, while 93.1% are public schools. The remaining share of 3% is covered by other educational institutions (Republic of South Africa, 2013).
In the *Education Statistics* of 2011, the DBE counted a total of 12.7 million students enrolled in General Education and Training, instructed by approximately 440,000 educa-

tors in about 31,000 schools in South Africa. About 11.8 million students are educated in public schools, while only 480,000 students visit an independent school. Whereas students in public schools are taught by 390,000 educators in 24,000 institutions, students in independent schools are taught by 31,000 educators in 1,500 institutions (Republic of South Africa, 2013). Howie & Sherman (2008) reported on the student-teacher ratio by ethnic groups of the South African population, stating that in 1989, the student-teacher ratio among White students was 17:1 with no teachers found to be under-qualified. On the contrary, the student-teacher ratio among Non-White students was 38:1, with almost twice as many students taught by one teacher compared to White students. Furthermore, over half of the teachers teaching Non-White students were considered under-qualified (Howie & Scherman, 2008). Some 20 years later, the student-teacher ratio in South African primary schools is still at 32:1, which is only slightly above the average student-teacher ratio of countries designated as having medium human development (Howie et al., 2018).

Language of Instruction

Somewhat unique to the educational system of South Africa are the principals on language of instruction. Guided by the SASA and manifested in the *Language in Education Policy* (LiEP) from 1997, use of native languages as language of instruction during pre-primary and early years of primary education, that is until grade 3, is ensured. However, from grade 4 onward, the LiEP requires English to be the primary language of instruction, and hence, 80% of students were taught in English in 2011 (Howie et al., 2012; van Staden et al., 2016). Additionally, the government requires students to learn a second language, that is either English or Afrikaans (Houserman, 2016). Again, the South African governments acknowledges 11 official languages, of which the most frequently spoken is IsiZulu (22.7%). In comparison, English is spoken by about 10% of the South African population (sec. 2.2.1). However, according to the DBE, the language of instruction in grade 3 is most frequently English (27.7%), followed by IsiZulu (20.1%), and IsiXhosa (14.0%) as of 2007. From grade 4 onward, 79% of students are taught in English, while those taught in another official language, for example, IsiZulu and IsiXhosa dropped to

1.5% and 3.1%, respectively, in 2007 (Republic of South Africa, 2012f).

School Enrollment & Completion Rates

Even though school enrollment increased since it became compulsory in 1996, Stats SA reported that in 2011, five million South Africans (27%) aged between 5 - 24 years did not attend formal schooling in an educational institution. 4.6% thereof are White, while 82.2% are Black-African. Looking at the Black-African population alone, 26.1% did not attend schooling in an educational institution. Although school enrollment of the Black-African population has increased over the past decade, respective results from 2011 show that about one third of the Black-African population in South Africa still lacks formal schooling (Republic of South Africa, 2012a).

Looking at the completion rates, Stats SA reported that in 2011, 8.6% of the South African population aged 20 years or older did not complete any education. Only 0.8% thereof were White, whereas 93.2% were Black-African. Looking at the Black-African population alone, only 10.5% did not complete any education. Among the White population, the greatest share of the population aged 20 years or older completed FET, that is grade 12. On the contrary, among the Black-African population of the same age group, the greatest share completed some secondary education (GET). Whereas 37% of White students aged 20 years or older completed higher education, only 9% of the Black-African students achieved alike (Republic of South Africa, 2012a).

Educational Disparities in South Africa

In 2012, South Africa showed, on average, 12.9 years of expected schooling, and 9.8 years of completed schooling. Numbers slightly increased according to latest data of 2019, resulting in 13.8 years of expected schooling, and 10.2 years of completed schooling. Overall, South Africa has come a long way since 1990, increasing mean years of schooling by 3.8 years on average (UNDP, n.d.). Furthermore, illiteracy rates in South Africa improved significantly during the last 35 years. Whereas in 1980, more than four million South Africans aged 15 years or older were incapable of reading or writing, this number

has been cut roughly in half, to about two million South Africans in 2015 (UNESCO, n.d.).

South Africa by International Comparison

The Economist published an article in 2017 that focused on the conditions of the educational sector in South Africa, claiming that even today, "South Africa has one of the world's worst education systems" (The Economist, 2017). The article built its arguments on latest empirical evidence collected in international large-scale assessments, for instance results of IEA's *Trends in Mathematics and Science Study* (TIMSS). First conducted in 1995, TIMSS evolved to be the longest running international large-scale mathematics and science assessment for fourth- and eighth-grade students. Performed on a quadrennial schedule, TIMSS provides individual performance scores that, all together, reflect the average performance of fourth- and eighth-grade students in each participating educational system. In 2015, TIMSS was administered in 60 educational systems around the world, among which highest results in mathematics were achieved in Singapore (618 score points); and lowest results were achieved in Kuwait (353 score points). Therefore, the range is 265 score points between the lowest and the highest ranking country. Despite its alleged improvements in the educational sector as mirrored by expected and mean years of schooling, South African fourth-grade students only achieved an average of 376 score points in the TIMSS 2015 mathematics assessment, which makes South Africa the second lowest-performing educational system within the pool of international participants. Results of eighth-grade students are similarly low. In mathematics, South African eighth-graders achieved an average of 372 score points, which is the second lowest performance score after Saudi Arabia (368 score points) (I. V. Mullis et al., 2015a). In the science assessment, South Africa participated with grade 8 only, achieving an average of 358 score points and coming in last on the international league table (I. V. Mullis et al., 2015b).
A similar pattern occurs when checking the latest available results of IEA's 2016 *Progress in International Reading and Literacy Study* (PIRLS), which has been testing student performance in reading since 2001. Of 61 educational systems participating in the study,

South African fourth-graders achieve, on average, 320 score points, which is the lowest result in international comparison. On the contrary, highest results were achieved in the Russian Federation, were students achieved, on average, 581 score points on the reading assessment (I. V. S. Mullis et al., 2017). Hence, results of PIRLS 2016 suggested much work needs to be done regarding the quality of literacy education in South Africa.

Factors fostering Educational Disparities in South Africa

Educational disparities in South Africa do not only occur externally and by international comparison, but rather first and foremost on a national level and along the lines of school and student characteristics. To identify causes of inequality in educational attainment and derive suitable solutions, emphasis has been put on schools and school resources in conjunction with teacher characteristics and teaching practices (Howie, 2003; Howie & Scherman, 2008; Howie et al., 2008; Ndimande, 2016; L. Zimmerman et al., 2011).

For instance, results from the TIMSS 1999 assessment showed that average science class achievement was about 130 score points higher in urban schools, and students taught by White teachers outperformed their counterparts in science by approximately 300 score points (Howie & Scherman, 2008). Furthermore, results from the PIRLS 2011 assessment showed that 42% of fourth-grade students received education in remote rural areas and scored, on average, over 100 score points less (428 score points) in reading than students educated in urban areas. On the contrary, 18% of students were educated in urban or suburban areas and scored, on average, around 530 score points on the reading assessment (Howie et al., 2012).

Furthermore, it was found that students who were taught in schools lacking resources for learning as well as a sufficient infrastructure achieved lower than students who benefit from well-equipped schools. For instance, lack of learning materials and school resources, high classroom size and instructional load, as well as lack of school management and governmental support have been found to be influential with regard to student performance in mathematics and science (Howie, 2003). Looking at shortage of school resources for

learning affecting teaching, students performed about 130 score points less in reading when being affected by such shortages. Distinguishing further by language groups, it became evident that students educated in native African languages were more likely to be affected by shortage of school resources than their peers. For example, 12% of fourth-grad students were greatly affected by shortages of school resources, whereas students educated in English did not experience any shortage of school resources (Howie et al., 2012). Moreover, Zimmerman et al. (2011) found that low-performing schools often lack effective strategies for curriculum improvement with regard to literacy development (L. Zimmerman et al., 2011). Ndimande (2016) argues, from a qualitative perspective, that South African educational policies are not sufficiently addressing respective issues, leading to the persistence of poorly equipped schools in rural areas. For example, township schools have been found to be affected by lower school quality, efficiency and effectiveness (Mestry, 2014), and hence, students being educated in townships and tested in English or Afrikaans performed lower than their peers being educated in remote rural areas (Howie et al., 2012). Furthermore, such schools are most often attended by socioeconomically disadvantaged students residing in rural areas of the country (Ndimande, 2016). On that note, van der Berg et al. (2011) stated that children of poorer, and hence, primarily Black-African neighborhoods, are especially prone to being educated in schools lacking sufficient resources and functionality, that is discipline as well as highly qualified teachers, and suffering from weak school management (van der Berg et al., 2011).

Against this background, the socioeconomic status of schools, represented by indices of school resources, was found to be highly predictive of student reading achievement as well. With an increase of schools' educational resources, student performance increased by about 69 score points. This is true even after controlling for other influential factors such as quality of teaching and time spent on reading tasks (van Staden & Howie, 2012).

In addition to the relevance of school characteristics, Howie (2003) also emphasized the relevance of student characteristics such as socioeconomic status, their self-concept of mathematical competencies, their attitudes towards the importance of mathematics, as

well as language spoken at home and language of instruction with regard to learning outcomes (Howie, 2003). Furthermore, student age and sex showed to be significantly predictive of reading performance. That is, girls perform better in reading than boys; and the older students in grade 5 are, the lower their average reading performance is (van Staden & Howie, 2012).

Governmental Efforts towards Improving Education

To counteract detected deficiencies in the South African school system reflected by official and international statistics, the South African government took different measures to achieving equity of educational opportunities for all.

For example, government expenditures on education decreased over the last decade, from 19.7% in 2012 to 18.4% in 2021 (UNESCO, n.d.). Hence, the educational sector receives the second highest financial expenditures of the South African government after general public services, which was allotted 24.4% in the fiscal year of 2103/2014 (Republic of South Africa, 2015). Stats SA reported that in 2013/2014, the largest shares of total government expenditures on education went to pre-primary and primary education (32%), followed by secondary and post-secondary education (29%), and finally, tertiary education, which was designated 24%. The remaining 15% of educational expenses fell into other education, which was not further classified (Republic of South Africa, 2015).

Another prominent example of governmental efforts to enhance equity of educational opportunities is the *Action Plan 2014: Towards Realisation of Schooling 2025*. Developed by the DBE as a milestone towards *Schooling 2025*, the Action Plan was introduced in 2010 aimed at ensuring the improvement of the quality of learning and teaching by establishing 27 national goals, at the center of which are the *three T's*. These are *textbooks*, *teachers*, and *time* (Howie et al., 2012; Republic of South Africa, 2012e). According to the Action Plan, the DBE set priorities for the educational targets of granting universal access to grade R for all children, developing and providing appropriate teaching and learning materials, and further pursuit of a standardized *Annual National Assessment* (ANA) among

grades 3, 6 and 9 to assess the quality of learning (Republic of South Africa, 2012e).[21] Given the fact that the majority of South Africa's population is under the age of 30 (sec. 2.2.1), the South African government further endeavored to increase the "ratio of young people who are in education, employment or training" by introducing pivotal programs, namely "professional, vocational, technical and academic learning programmes, which meet critical needs for economic growth and social development" (Republic of South Africa, 2012e, p. 156). Counteracting a severe lack of infrastructure in South African schools, the government furthermore attempted to eliminate insufficient school infrastructures and ensure water and electricity supply as well as basic sanitation to more than 1,000 schools after 2013 (Motshekga, 2012).

Apart from measures aimed at improving the quality of learning and teaching as well as increasing participation rates in tertiary education, the South African government further sought to foster the accessibility to education. In line with such attempts, the SASA of 1996 permitted public schools to raise private funds in order to supplement governmental expenditures in the educational sector (Republic of South Africa, 1996). Consequently, many schools in the public sector introduced school fees which allowed them to procure resources for learning, school equipment, and facilities as well as extend resources for personnel (Republic of South Africa, 2006). However, respective fees are partly determined by "school governing bodies based in part on what the families of the pupils in these schools can afford to pay [resulting in] only limited racial mixing in schools" (Ardington et al., 2011, p. 530). Though beneficial on some level, such prevailing policies first and foremost enhance the privatization of public schooling, and hence, reinforce racial segregation of the poorer and mainly Non-White population (Ardington et al., 2011; Motala & Sayeed, 2009).
On that note, the DBE stated in 2006 that South Africa's poor primarily residing in the former homeland areas

[21] Initially introduced in 2010, ANA collects data from a written assessment administered among all three- and sixth-graders as well as a sample of ninth-graders (Howie et al., 2012).

> "contributed a disproportionate share of their incomes over many decades to the building, upkeep, and improvement of schools, through school funds and other contributions, including physical labour [*sic*]. (...) Such contradictions within the same public school system reflect past discriminatory investments in schooling [during Apartheid]." (Republic of South Africa, 2006, p. 10)

Thus, originally implemented in order to ensure an "education of progressively high quality for all *learners*" (Republic of South Africa, 1996, p. 5), prevailing school fee policies perpetuated remaining disparities in the public-school sector between the socioeconomic advantaged and disadvantaged of the South African population rather than mitigating them.

Aiming to counteract the emerging trend of public school privatization in South Africa and counteract inequalities of educational opportunities, the *Education Laws Amendment Act* of 2005 was introduced to set the grounds for no-fee schools (Mestry, 2014; Republic of South Africa, 2012e). According to the *Amended National Norms and Standards for School Funding* (ANNSSF) introduced by the DBE in 2006, no-fee schools qualify for increased allocations of governmental financial aid for non-personnel, non-capital expenditures (Republic of South Africa, 2006; van der Berg et al., 2011). Respective resources can be used to cover costs such as maintenance and service expenses, as well as learning materials (Motala & Sayeed, 2009). Hence, the no-fee school policy is considered to be a pro-poor policy (van der Berg et al., 2011, 2016). In 2010, the South African government reported that more than 80% of South African public schools, that is the majority of school in the country, were identified as no-fee schools (Republic of South Africa, 2012e). However, despite good intentions, no-fee school policy cannot be considered the remedy to the dysfunctionalities in the South African education system (Spaull, 2019). Rather, socioeconomically low-performing students often remain in socioeconomically disadvantaged schools, and hence, "the probability [for these students] of 'succeeding against the odds' when attending one of these schools is dismal" (Spaull, 2019, p. 3). Furthermore, the no-fee school policy was originally planned to be applied to compulsory education

only, that is grade 1 - 9 (Republic of South Africa, 2006), thus, leaving parents to provide resources for any post-secondary and tertiary education themselves. Although subsidized by governmental funding, tertiary education can be fairly expensive. Depending on the university and the aspired degree, tuition fees can vary between ZAR30,000 - ZAR60,000 ($1,700 - $3,500) per year. Although Branson et al. (2015) argue that additional grades benefited from prevailing policies as well (Branson, Hofmeyer, & Lam, 2012), the unequal distribution of net wealth among ethnic groups (sec. 2.2.3) points to the assumption that tuition fees can be prohibitive for poorer communities, and thus, discourage participation in further education all together.

To address these concerns, the South African government allows for the exemption of school fees. As secured in the ANNSSF of 2006, parents who are unable to pay the required school fees can apply for exemption accordingly.[22] Aimed at ensuring that education is accessible to all, the exemption of school fees also raised several critiques, including Roithmayer (2003), who argued that respective policies actually limit accessibility to education rather than foster it. This is because, first, many families decide not to apply for exemption due to the time-consuming process of gathering information and the application procedures as well as the fact that many schools are discriminating against those who receive exemptions. As a result, school fee exemptions are, in fact, rarely applied for or received. Second, school fee exemption does not cover additional costs such as transportation, uniforms, stationary and the like.[23] Third, governing bodies of educational institutions evidently fail to inform parents of their eligibility to receive exemptions or deny applications (Roithmayer, 2017).

Hence, despite many attempts to achieve equal access to quality schooling through pro-poor policies, school fees have not been disestablished by the South African government

[22] Exemptions are furthermore granted to orphans, abandoned children, as well as recipients of a poverty-linked state social grant (Motala & Sayeed, 2009; Republic of South Africa, 2006).

[23] The governing bodies of respective educational institutions may decide independently what is included to school fees. If certain expenses such as transportation and uniforms are considered separate, governing bodies may excluded alike from exemptions, which is when parents are required to cover such costs regardless of their income (Roithmayer, 2017).

yet, and thus, quality education remains elusive for a majority of South Africa's poor (van der Berg et al., 2011). That is especially true for Black-African students, who are primarily educated in dysfunctional schools that fail to ensure sufficient educational attainment in both numeracy and literacy performance (Spaull, 2013). Hence, while access to quality education was constrained by ethnic belonging during the Apartheid Regime, it is now determined by class and socioeconomic status (Spaull, 2019).

2.3. Summary and Implications for the Study

Aimed at emphasizing the characteristics of South Africa's socioeconomic low-performers, a rough overview of South Africa's political, social and economical situation was presented. Considering the importance of South Africa's political past for the country's present state, first, relevant milestones of political events depicting the rise and fall of the Apartheid Regime were presented. Second, segregation policies during Apartheid were summarized, briefly introducing two major aspects, that are residential and educational segregation. Lastly, the present state of South Africa was conceptualized using three dimensions of the HDI, which are health and life expectancy, standard of living, and education. Additionally, a short overview of South Africa's demographic facts was provided. The following is summarized from this section.

Introduced by Desmond Tutu and adapted by Nelson Mandela as South Africa's first democratically elected president over 20 years ago, the picture of the Rainbow Nation still fits the profile of South Africa's population. However, the originally envisioned melting pot of colors was superseded by a co-existence of various shades, that are ethnicities, languages, and beliefs, capturing the repercussions of Apartheid that South Africa has not yet overcome (Haferburg & Osmanovich, 2017; van der Berg et al., 2011). The Apartheid Regime and its respective policies of racial segregation split the country in favor of the White population. Fostering white supremacy, the Non-White population experienced discrimination due to skin color, cultural belonging, and social status, altogether limiting opportunities of social mobility and fostering a two-tier society. Reviewing official na-

tional and international statistics of the past decade, it became apparent that respective indicators which separated South Africa in the past still play an important role for the constitution of its present societal patterns.

For instance, poverty levels are exceptionally high among the Black-African population. In 2015, more than half of South Africa's poor lived in remote rural areas, such as townships and former homelands, and were affected by overpopulated living conditions. Furthermore, drawing on results from the 2011 PIRLS assessment, students educated in a native African language are more likely to be affected by shortages of school resources, and fourth-grade students receiving education in remote rural areas performed lower when tested in English or Afrikaans compared to their peers. Despite the fact that policies of educational segregation are counteracted by governmental attempts of pro-poor policies and education for all, almost a third of the South African population did not attend schooling in 2011, about 80% of which were Black-Africans (sec. 2.2.4). Thus, Black-Africans with no educational degree have significantly lower chances of accessing a well-paid career, and to achieve sufficient income as well as life opportunities.

To conclude, the following conceptualization of socioeconomic high- and low-performers is deduced from the information presented in this section.

First, the group of socioeconomic high-performers primarily consists of the White population. This group has more access to monetary resources and preferable working environments, benefits more often from better health and living conditions, primarily speaks English/Afrikaans, is located in the urban and suburban areas of the country, and holds higher educational degrees. Second, the group of socioeconomic low-performers primarily consists of the Non-White, mostly Black-African population. This group has access to fewer monetary resources and less preferable working environments, suffers more often from poor health and living conditions, primarily speaks a native African language, is located in the remote rural areas of the country, and holds lower educational degrees (Burger et al., 2015). The group of socioeconomic low-performers is the focus group of this study, and thus, serves as the reference group for the empirical work of this study.

3. Theoretical Framework

The theoretical framework of this study is based on the *Rubicon Model of Action Phases* introduced by Heckhausen & Gollwitzer (1987). Originating from the field of psychology, the Rubicon Model was designed to reflect the entirety of actional behavior, incorporating motivational, volitional and evaluative phases as well as the sequential order thereof. In contrast to other concepts solemnly focusing on either the choice of action goals, that is the motivation for an action, or the realization thereof (J. Heckhausen & Heckhausen, 2008), the Rubicon Model of Action Phases achieves a combination of both aspects and thereby provides an exhaustive framework of human action behavior. Hence, the Rubicon Model is applicable to diverse contextual settings; for example, to systematically explore the field of work motivation and action behavior in the work environment (Nerdinger, 1995).

The generalizability of the framework is particularly conducive to this study since it allows for the incorporation of both focal points, that is the qualitative exploration of general attitudes towards educational investment behavior among socioeconomic low-performers in South Africa; and the quantitative exploration of parental involvement as a form of educational investment, and with regard to student outcomes in learning among socioeconomic low-performing families in South Africa.

To that end, this section presents the Rubicon Model of Action Phases first. Next, additional theoretical concepts are introduced to substantiate single phases of action behavior acknowledged by the Rubicon Model. Respective concepts originate from *Rational Choice Theory*, and focus on general rational decision making. Lastly, theories of parental involvement and the *Model of Parental Involvement Process* are introduced.

3.1. Approaching Action Behavior

3.1.1. The Rubicon Model of Action Phases

Throughout the last century, the psychological landscape of exploring human motivation was dominated by concepts of *Expectancy-Value Theory*, focusing on *pre-decisional* as-

pects of actional behavior (e.g., Atkinson, 1964). In the late 1960s, Locke (1968) expanded beyond existing concepts and provided empirical evidence that pointed towards the importance of *post-decisional* aspects as well (Locke, 1968). Assuming the inseparability of motivation and action, and the distinction between different phases of human actional behavior, Heckhausen & Gollwitzer (1987) finally came to the conclusion that the pre-decisional and post-decisional phenomena could successfully be united in an exhaustive framework of human actional behavior (H. Heckhausen & Gollwitzer, 1987).
Building on the work of Kuhl (1983) and Heckhausen & Kuhl (1985), who argued that motivation and action function naturally different from each other (H. Heckhausen & Kuhl, 1985; Kuhl, 1983), Heckhausen & Gollwitzer (1987) proposed that the pre-decisional phase equals *motivation*, and the post-decisional phase refers to *volition*, that is *action*.

> "The motivational state of mind is terminated with the making of a decision, a more or less conscious event that launches the individual into the volitional state of mind [which] entails consideration of when and how to act for the purpose of implementing the intended course of action." (H. Heckhausen & Gollwitzer, 1987, p. 103)

Based on empirical evidence drawn from a series of psychological experiments testing the motivational and volitional state of mind, the Rubicon Model of Action Phases was finally developed. Inspired by Julius Cesar, who crossed the Rubicon and thereby irreversibly initiated civil war, Heckhausen & Gollwitzer (1987) used this picture and named the transition from contemplating to enacting an intent "a psychological Rubicon, a boundary line between different states of mind" (H. Heckhausen & Gollwitzer, 1987, p. 120). By establishing the Rubicon Model, the incorporation of different stages of actional behavior while simultaneously expatiating their separation from each other was achieved. Thereby, the Rubicon Model unifies the two focal points of motivational psychology, namely the selection of action goals and their execution. The model is therefore treated as the most exhaustive framework of human action behavior at present (Kirchler & Walenta, 2010). Figure 1 shows the Rubicon Model of Action Phases.

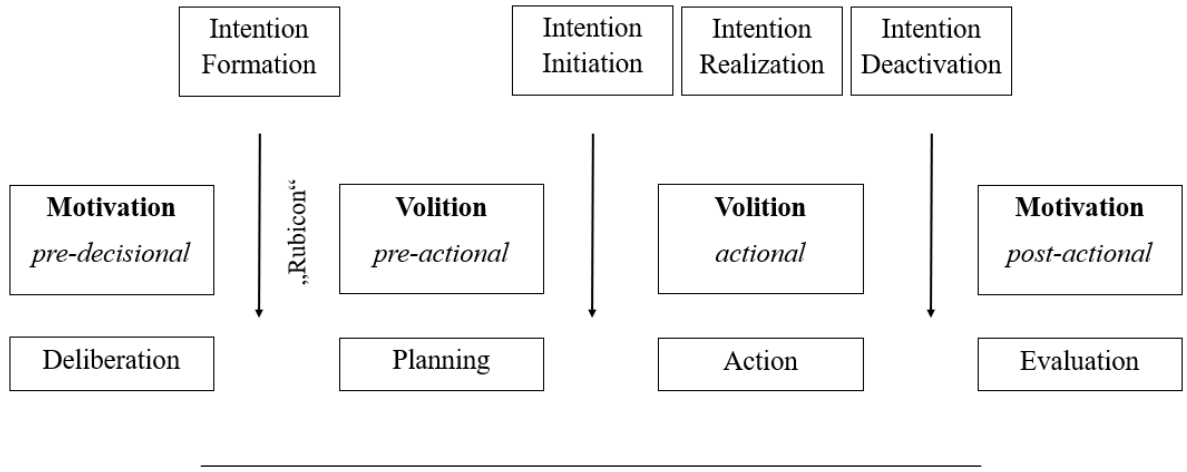

Figure 1: The Rubicon Model of Action Phases, (after J. Heckhausen & Heckhausen, 2008).

The Rubicon Model distinguishes between two dimensions, namely motivation and volition. Motivational processes focus on choosing and setting goal intentions [Zielintentionen]. Phases referring to motivation are therefore phases of *goal setting* (Gollwitzer, 1990). On the contrary, volitional processes aim at achieving a goal intention. Different from motivation, volition thereby refers to *goal striving* (Lewin, 1926).
Furthermore, the model is separated into four phases assigned to one of the two dimensions, namely *motivation* and *volition*. These are the *pre-decisional*, *pre-actional*, *actional*, and *post-actional* phases. All phases are accompanied by corresponding activities, which are *deliberating* [Wählen, Abwägen], *planning* [Planen, Zielsetzung], *acting* [Handeln] and *evaluating* [Bewerten]. The model is understood as a linear progression of events with each phase leading to the next, finally resulting in the evaluation of action outcomes. Along those lines, experiences drawn from the process of actional behavior are likely to influence future decision making (J. Heckhausen & Heckhausen, 2008). An in-depth description of the respective phases defined in the Rubicon Model are presented below.

The first phase is known as the pre-decisional phase of motivation and focuses the choice of a specific action from a set of possibilities evolving from one's individual wishes, motives and needs. As stated above, this phase is allocated to goal setting. In the pre-decisional phase, individuals assess both the desirability and feasibility of different goals. During

that process, an individual is confronted with different questions relevant to goal setting. Those are, for example, whether certain resources, such as time, are available to realize a goal, whether positive or negative consequences are to be expected, and whether these are likely to occur. Finally, individuals balance the desirability of expected outcomes with the feasibility of achieving them, resulting in a binding decision on a goal intention, that one is likely to aim at achieving. After the decision is taken, a feeling of commitment emerges, which deters individuals from considering quitting the execution of a certain goal intention (J. Heckhausen & Heckhausen, 2008). This emphasizes the importance of the pre-decisional phase, given that it adds to the determination of an individual to successfully turn a decision into an action.

The second phase is known as the pre-actional phase of volition and focuses on the gap between setting a goal intention and achieving it. Logically, this phase is assigned to volition, which indicates that the process of deciding on a goal intention has been terminated, and is now followed by the attempt to realize it. In turn, wanting to achieve a goal intention becomes the major motivation for further actions. However, in some cases, a goal cannot be reached immediately but requires time, for example the completion of educational degree (Nerdinger, 1995). Individuals in the second phase of the Rubicon Model must wait for a suitable opportunity in order to reach their goals long-term. To pursue the achievement of a goal intention, plans and resolutions commonly identified as *implementation intentions* [Durchführungsintentionen] are made. Respective resolutions often entail planning of when and how goal-oriented actions can be taken, for example scheduling exams at university towards completing a desired educational degree. Finally, resolutions support the management of difficulties occurring while aiming at successfully achieving a particular goal (J. Heckhausen & Heckhausen, 2008).

Seamlessly following the pre-actional phase is the third phase, known as the actional phase of volition. It focuses on the execution of the resolutions and action plans, that one decided on during the first and second phase of actional behavior.[24] Hence, within phase

[24] Nerdinger (1995) summarizes different theoretical approaches stemming from the field of organizational psychology that underpin the phase of volition, for example the *Law of Determination* after

three, one acts on achieving a goal intention that has been set and planned during phases one and two. Whereas the pre-decisional phase of the Rubicon Model is allocated to goal setting, the pre-actional and actional phases are dedicated to goal striving. After goals have been set, and plans to achieve these goals have been made, individuals then attempt to successfully carry out respective plans or finalize necessary actions (J. Heckhausen & Heckhausen, 2008).

The fourth phase is known as the post-actional phase, and focuses on the evaluation of results with regard to the goal intention set in the first phase. Typical questions that one might address in order to evaluate the results of an action are whether the desired outcome was achieved, the action was successfully completed, the process of achieving a goal intention was easy or difficult, and, most importantly, whether additional steps are possible in case a goal intention has not been reached yet. If a goal intention is reached successfully, however, it is deactivated, and the process of actional behavior is completed (J. Heckhausen & Heckhausen, 2008).

Although widely acknowledged, criticism of the Rubicon Model has been noted as well. For example, Nerdinger (1995) argued that the model does not emphasize external situational constraints [situative Zwänge] (Nerdinger, 1995). However, Nerdinger (1995) among others applied the Rubicon Model to explain processes of motivation and action within special circumstances of the labor market, where the the absence of external goal intentions might be accompanied by penalties from others, for example the employer. This is not the case for the subject of interest in this study, which is the mechanism of general educational investment behavior in one's own educational career as well as parental involvement with regard to student outcomes in learning. Both actions are considered voluntary, and sanctions for non-completion are not expected. Hence, the Rubicon Model of Action Phases is still considered a good fit for the fundamental framework of this study, although it does not account for external situational constraints.

Ach (1935), and Locke & Latham (1984), as well as the *Theory of Action Control* after Kuhl (1983), and the *Theory of Self-Regulation* after Karoly (1993) (Nerdinger, 1995).

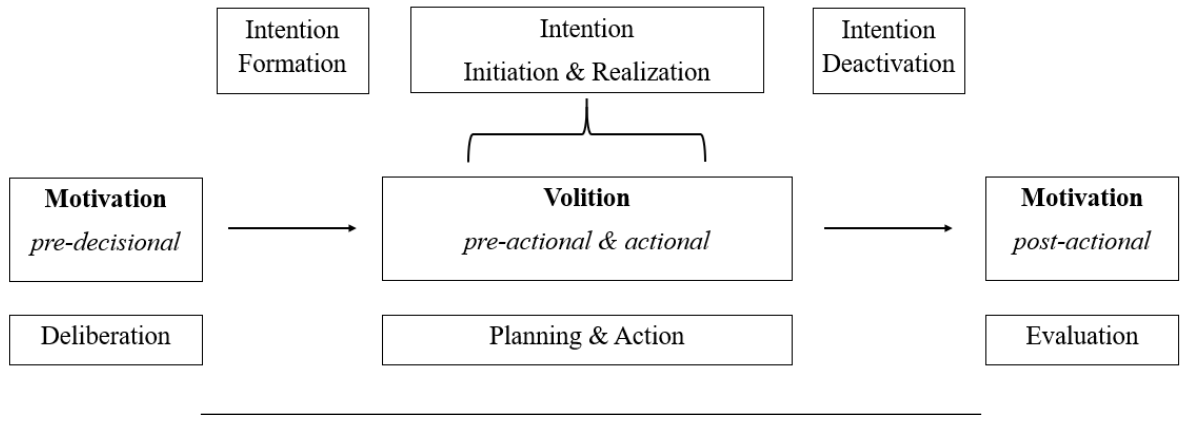

Figure 2: The Rubicon Model of Action Phases. Adapted figure, based on the original model (after J. Heckhausen & Heckhausen, 2008).

3.1.2. Implications for the Study

In summary, the Rubicon Model provides an exhaustive framework depicting the process of action motivation, volition and the evaluation thereof. Thereby, it emphasizes relevant elements of the process of actional behavior and points towards the linear progression of events.

In order to suit the needs of this study, the Rubicon Model is slightly adapted. While phase one (pre-decisional motivation) is operationalized as such, phase two (pre-actional volition) and phase three (actional volition) are joined for the purpose of this study. This is because a distinction between pre-actional and actional behavior cannot be made with the data sets utilized for this study. On that note, Kuhl (1983, 1984, 1987) and Nerdinger (1995) support this approach, stating that the difference between the motivational and volitional phase is the core distinction when explaining the process of human action behavior, and further distinctions of the volitional phase are, therefore, left unattended. Finally, this study emphasizes the sequential character of the Rubicon Model, depicting the linear progression of its phases by introducing arrows to the model. Figure 2 shows the adapted version of the Rubicon Model accordingly.

Although exhaustive in its conceptualization, the Rubicon Model is a generic theoretical model. Hence, it does not aim to provide in-depth information about contextual specifics of its phases. Furthermore, the Rubicon Model does not account for differences in actional

behavior as a function of individual backgrounds, such as socioeconomic status. Therefore, the following sections focus on substantiating the Rubicon Model by the focal aspects of this study. These are, first, rational decision making in education and, second, specific educational investment behavior represented by parental involvement. The aspect of socioeconomic status is considered a predictor for respective mechanisms as well.

3.2. Motivation in the Context of Rational Decision Making

3.2.1. Core Needs Motivating Actions

The pre-decisional phase of the Rubicon Model involves the decision for a certain action from a set of possibilities and is referred to as motivation. Robbins & Judge (2017) define motivation "as the processes that account for an individual's intensity, direction, and persistence of effort toward attaining a goal" (Robbins & Judge, 2017, p. 247). Hence, the motivation of a person is associated with the goal one is aiming to achieve, and the choice of an action is not a random decision, but determined by certain needs, namely motives. The present literature holds various general approaches aimed at identifying the core needs motivating human action. Three widely acknowledged concepts of general motivational theory are presented by Maslow (1943, 1954), Alderfer (1969), and Esser (1999).

In the 1940s, Maslow developed the Hierarchy of Needs, a model originally consisting of five different levels in a hierarchical order (Maslow, 1970).[25]
The first level refers to physiological needs such as hunger, thirst, shelter, and sexual contact. After those are deactivated due to their saturation, second-level needs referring to safety-security and focusing on the protection from physical and emotional harm become relevant. The third level summarizes needs of social-belonging and attention. Those are affection, sense of belonging, acceptance, and friendship. The fourth level focuses internal

[25] Recently, the inclusion of a sixth level on intrinsic values has been proposed. Robbins & Judge (2017) argue though, that the validity of a sixth category has to gain acceptance before being accounted for as part of the model (Robbins & Judge, 2017). Therefore, only the original five levels are included here.

and external factors ensuring self-esteem. Internal factors are self-respect, autonomy, and achievement; external factors are status, recognition, and attention. The fifth level finally refers to self-actualization and represents the desire to improve. Concepts of growth, realizing individual potential, and self-fulfillment become relevant (Robbins & Judge, 2017). Thus, saturating needs of level five enhances personal development. Different from needs at level one to four, which are categorized as needs activated by deficits, needs of level five can never be completely satisfied (Kirchler & Walenta, 2010).
Generally, deficits call attention to respective needs, and their satisfaction forces the deactivation of need-fulfilling activities. Hence, the activation of a higher-level need is possible only if the needs at preceding levels have been satisfied, and hence, deactivated (Robbins & Judge, 2017). Even though Maslow's theory is widely acknowledged in the field of organizational psychology due to its innovative character regarding self-activation, researchers had difficulties finding empirical evidence for the model's validity. Especially the core idea of a hierarchy of needs is yet not verified (Kirchler & Walenta, 2010).

Trying to address the issue related to Maslow's model, Alderfer (1969) revised the original set of needs and proposed three groups of core needs instead. Those are *existence*, *relatedness*, and *growth*. *Existences* summarizes the first two levels of Maslow's model, and subsumes basic material needs. *Relatedness* concerns interpersonal relationships and is equivalent to Maslow's third and fourth level of needs. Finally, *growth* represents needs of personal development and is in line with the fifth level of the original model, namely self-actualization (Alderfer, 1969). Different from Maslow, Alderfer (1969) did not claim for the model to be hierarchical. Instead, he argued for the possibility of satisfying more than one need at a time. Alderfer (1969) thereby opened the model to different cultures and preferences of personal needs, respectively (Kirchler & Walenta, 2010).

Originating from the field of sociology and referring back to Adam Smith, Esser (1999) reduced the scope of human needs even further and introduced the *Social Production Function*. According to Esser, human action is motivated by the constant need to gen-

erate utility (Esser, 1999). In order to do so, corresponding needs must be met. Esser defined two general needs that he identified to be conditional for the production of utility. These are *social appreciation* [Soziale Wertschätzung] and *physical well-being* [Physisches Wohlbefinden] (Esser, 1999).

In order to saturate those needs, investments are taken in the form of production factors such as properties, resources, objects, goods and services. Esser (1999) defined such factors as (primary) intermediate goods [Zwischengüter] (Esser, 1999). Intermediate goods are subjected to two premises. First, they are determined by the social context a respective action is associated with. Second, intermediate goods are only in the rarest of cases directly accessible to the actor, which is why they need to be produced before they can be used for the satisfaction of social appreciation and physical well-being. Respective elements which are needed for the production of intermediate goods are indirect intermediate goods. Esser (1999) identified the factor of time as always being decisively involved in the production of respective goods (Esser, 1999). In contrast to primary intermediate goods, which are culturally constructed, indirect intermediate goods are subjected to an objective scarcity, which is not externally influenced, for example by changes of social norms and standards.

Finally, Social Production Functions can be connected to a causal chain, resulting in the production of utility: using time, indirect intermediate goods are produced, which are elementary for the production of primary intermediate goods. Ultimately, (primary) intermediate goods serve to produce the needs of social appreciation and physical well-being, and thus, lead to the production of utility (Esser, 1999).

In summary, across all three presented concepts of core needs motivating human action, the saturation of a certain need functions as a motive for actional behavior. Although concepts differ in the extensiveness of their description and structure of their hierarchies, they generally refer to two core motives for action, namely physical and psychological well-being. Of particular relevance is the element of time, which is assumed to indirectly influence the achievement of saturating needs.

Following this introduction to core needs motivating human action behavior, the question arises according to which paradigm individuals decide to act or not act on a certain need, namely a motive. For example, in the case of hunger the answer is obvious: apart from disruptive elements, one most likely will aim at satisfying the need of feeling hungry. But based on which parameters does one act on a need of growth, for example in terms of educational attainment? And by what means does one choose certain alternatives over others? In order to answer respective questions from a theoretical perspective, theories of rational decision making and, more precisely, Expectancy-Value Theory, proved to be beneficial (Nerdinger, 1995).

3.2.2. Expectancy-Value Theory as a Paradigm for Acting on a Motivation

As mentioned above, Expectancy-Value Theory (EV-Theory) [Erwartungs-Wert Theorie; Wert-Erwartungstheorie] has guided the scientific approach to explaining human action behavior from the angle of rational decision making for a long time. Based on the thoughts of Daniel Bernoulli, who established the basic idea of EV-Theory in economics, the paradigm also proved to be beneficial for other scientific areas, first and foremost in the field of psychological theory of action. Kurt Lewin (1944) contributed significantly to the development of EV-Theory in psychology, followed by Atkinson (1957), and Heckhausen (1977). A sociological yet economically motivated concept of EV-Theory is presented by Esser (1999) who postulated a generic framework to systematically explain rational human decision making. Referring to its premises, Esser states that EV-Theory is the most generic answer to the question why someone acts in a particular way (Esser, 1999).
In line with Becker (2000), this study chooses Esser's approach to EV-Theory for two reasons. First, it incorporates existing and widely acknowledged approaches on EV-Theory (e.g., Boudon, 1974; Breen & Goldthorpe, 1997; Erikson & Jonsson, 1996); second, aimed at finding an overarching yet exhaustive framework for the existing body of approaches on EV-Theory, the conceptualization of Esser (1999) is less complex and therefore straightforward to apply (R. Becker, 2000).
Moreover, EV-Theory has been studied most extensively with regard to its applicability

for educational decision making (Stocké, 2010). This is beneficial for this study as the existing body of research (sec. 4) on the applicability of the theory offers a useful point of reference for a theory-based interpretation of educational investment behavior in South Africa.

The idea of EV-Theory is straightforward. As the name implies, it consists of two core elements, that are expectations and the valuation thereof. On that note, EV-Theory is based on cost-benefit calculations (R. Becker, 2000). That is, actional behavior along the lines of EV-Theory entails weighing the value of an action alternative against its probability to be successful. Yet action alternatives that entail high costs without appearing to be beneficial are untended. To that end, EV-Theory follows the logic of selectivity (Esser, 1999, p. 249). Esser (1999) installed six essential premises of an economic-based approach to expectancy-value guided decision making (Esser, 1999).

First, Esser (1999) argued that every action is based on a selection of possible action alternatives. In order to explain the selection of an action, all possible alternatives must be determined accordingly. EV-Theory assumes that there are always at least two action alternatives, which are mutually exclusive. The pool of possible action alternatives is referred to as the space of alternatives [Alternativenraum], denoted by A_i (Esser, 1999). Second, the results of selective action behavior are outcomes, denoted by O_j. Third, the outcomes of each action alternative are valued regarding their utility. That is, an actor might find the utility of an outcome to be somewhat positive, negative, or even neutral. The valuations of action outcomes are most likely subjective evaluations, deviating from any objective value. Along those lines, expected positive outcomes are referred to as utility (U), whereas expected negative outcomes are considered as costs (Esser, 1999). Fourth, the actor considers the probability for an outcome to eventuate. Those probabilities are the actor's expectations, denoted by p_{ij}. Esser (1999) further subdivided the expectations p_{ij} for the evaluated outcomes U_j into four cases, namely security, risk, uncertainty, and ambiguity (Esser, 1999). The estimated mean of those cases forms the starting point for the expectations p_{ji}. Statistically, the higher the knowledge about the probability of

achieving a certain outcome, the lower the scatter around the starting point. In the case of security and risk, the variance is minimal; but in the case of uncertainty, it reaches its maximum (Esser, 1999). Fifth, action alternatives are weighted through the subjective expected value or expected utility of an action, denoted by EU-weights. According to Esser (1999), the evaluation of the alternatives of actions must be understood as "the core of the logic of selection according to the expectancy-value theory [der Kern der Logik der Selektion nach der WE-Theorie]" (Esser, 1999, p. 256). Sixth, actors finally select an action alternative based on the evaluation of possible outcomes, considering the valued utility and the expected probability for the eventuation thereof. Hence, the action alternative with the best possible and expected utility is most likely to be chosen (Esser, 1999, p. 248). "The rule of the logic of the selection of action is accordingly the *maximization* of the expected utility [Die Regel der Logik der Selektion des Handelns ist demnach die *Maximierung* des erwarteten Nutzens]" (Esser, 1999, p. 258).
Finally, Esser (1999) pointed towards the causality of the premises of action behavior implied by EV-Theory. Given the consideration of consequences of action alternatives, the individual evaluation thereof, and the probability that respective consequences occur, the choice of an action alternative is understood as a causally progressing chain of events (Esser, 1999).

In summary, the paradigm of EV-Theory implies that humans act against the background of certain goal intentions, as well as the evaluation of the probability to what extent these intentions are likely to be realized. On that note, the weight of subjective expected utility (EU-weight) is decisive for the choice of an action alternative. Referring back to the aforementioned question, individuals are likely to choose action alternatives from a set of motives along the lines of cost-benefit calculations as addressed within EV-Theory. That is, they are likely to chose an action alternative that accommodates their needs while being minimally cost-intensive and promising the best chances of success.

3.2.3. Expectancy-Value Theory as a Rationale for Educational Investments

The supplementary question resulting from general concepts of rational decision making and EV-Theory, respectively, is whether they can serve as a rationale for the explanation of educational decision making. On that note, several reflections on the suitability of such concepts have been made, collectively aiming at explaining the mechanisms of educational decision making, and persisting educational inequalities as a result thereof. In conjunction therewith, socioeconomic resources have been identified as a main driver.

Early work by Eccles et al. (1983) proposed a model of educational decision making based on EV-Theory, which assumed that achievement-related educational choices are directly associated with expectancies and values thereof. For example, an educational choice ('I study mathematics') is made considering the utility value of that choice ('I pass my mathematics class successfully'), as well as the required costs ('I cannot play a game while studying'). Naturally, academic performance in domains such as mathematics, science or literacy is subjected to choosing beneficial educational action alternatives, for example to study a subject in order to achieve a goal such as performing well on a test (Wigfield & Eccles, 2000).

On that note, socioeconomic status has been considered as a decisive factor in the equation as well. For example, Erikson & Jonsson (1996) tied in with the work of Boudon (1974) and argued that students of lower social backgrounds are cognitively disadvantaged compared to their peers. Furthermore, the gap of educational outcomes between students with lower and higher social backgrounds increases with the educational status of the parents as well as their ability to support their children's schooling. Hence, educational success of students is likely to be positively associated with the availability of economic, cultural, and social capital of the family environment (R. Becker, 2000).

Finally, Esser (1999) added to the work of Erikson & Jonsson (1996) and stipulated an EV-Theory guided framework for educational decision making. Utilizing the example of points of transition, that are time points of students transferring from primary to sec-

ondary school,[26] Esser (1999) aimed at explaining the mechanisms of parental educational choices for their children. Thereby, theoretical reflections are based on one's subjective expectation of utility, including rational thinking about costs, utilities, and expected probabilities of success (R. Becker, 2000, p. 454). According to Esser (1999), the general logic of EV-Theory applies to the particular case of educational decision making at points of transition. For instance, aiming to decide on an educational pathway, students and parents have two action alternatives to choose from. They are, to either climb the educational ladder and continue in a higher educational track (A_b), or to not do that, and thus, remain in a lower educational track (A_n). The choice of an action alternative is accompanied by the valuation of three outcomes. First, the value of an outcome's utility (U) is considered. Second, the costs (C) of an educational decision are accounted for. Among those are both direct expanses as well as opportunity costs, which emerge from income loss while obtaining an educational degree. Third, the fear of loss of status (SV) is decisive for the choice of an action alternative, that is either continuing in a higher or lower school track. Finally, respective action alternatives are accompanied by two expectations. These are, first, the expectation (p) of realizing the value of an outcomes' utility (U); and second, the expectation (c) of risking loss of status (-SV) when forgoing an educational opportunity (Esser, 1999). The expectations (p; c) are appointed to be parameters of the probability for a certain action alternative to be chosen (R. Becker, 2000).

Based thereon, Esser (1999) presented the following equations, expressing the expected value of the action alternatives in educational investments, which are (1) refraining from continuing in a higher school track, or (2) proceeding with it (Esser, 1999):

$$(1)\ EU(A_n) = c\ (\text{-}SV)$$

$$(2)\ EU(A_b) = pU + (1 - p)\ c\ (\text{-}SV) - C > c\ (\text{-}SV) = U + cSV > C/p$$

Furthermore, Esser (1999) summarized that the expected utility of an educational investment needs to be higher than the expected utility of a non-investment in order for parents

[26] Effects of socioeconomically driven educational decision making at points of transition have been identified as *secondary effects* (Boudon, 1974; Erikson & Jonsson, 1996; Maaz et al., 2008). However, respective theoretical approach is irrelevant to this study, and hence, untended here.

to decide on an educational investment, that is, for example, the choice of a higher school track. Hence, the following is true:

$$EU(A_b) > EU(A_n)$$

In addition to the traditional motives of rational decision making, namely educational returns such as income, job safety, and a prestigious profession (Stocké, 2010), Esser (1999) identified fear of loss of status (SV) or the need of status preservation, respectively, to be of relevance (Esser, 1999).

The point of reference for status preservation is the socioeconomic status of the family, determining whether an educational decision leads to upward or downward mobility of the next generation. Furthermore, the prevention of downward mobility functions as a stronger motive for educational investments than the achievement of upward mobility (Stocké, 2010). Considering the stated parameters, Esser (1999) reasoned that the higher the probability for educational success and the need for status preservation is rated, the higher the motivation to make an educational investment. Consequently, parents of lower socioeconomic status do not fear social descent of their children as much as their counterparts, given that status preservation is less relevant to them. Furthermore, the risk of educational investments is bigger for parents with fewer resources to draw on, given that educational success might be less likely. This is not least due to missing social and cultural capital, such as the ability to support the child effectively throughout formal schooling (Esser, 1999; Stocké, 2010). Hence, socioeconomic status is decisive to the evaluation of utility or cost-benefit calculations, given that parents with access to more resources have better chances to compensate for unexpected deficiencies in educational performance that might jeopardize the overall goal of status preservation. For instance, parents of higher socioeconomic status are more likely to pay for private education or tutoring in order to support their children in achieving a certain educational goal. Furthermore, parents of higher socioeconomic status are more likely to have the necessary educational status themselves to support their children in troubleshooting learning deficiencies (R. Becker, 2000; Esser, 1999; Stocké, 2010).

To conclude, the subjective expectation of the utility of action alternatives forms the motives for rational decision making in education. Along those lines, presented concepts and frameworks emphasize the decisiveness of socioeconomic status for the evaluation of educational costs and benefits, and the success of educational investments respectively. Consequently, children with families with lower social backgrounds are less likely to get the chance to excel academically, whereas their counterparts experience the opposite, and are even cushioned by their families' resources in the case that the expected educational success fails to materialize. Respective mechanisms commonly lead to an inheritance of family capital, that is "the apple doesn't fall far from the tree" (Esser, 1999, p. 265). Whereas this is advantageous for socioeconomically high-performing families, it also entails diminished life opportunities for socioeconomically low-performing families.

3.2.4. Implications for the Study

By introducing the Social Production Function, Esser incorporated the core concepts proposed by Maslow (1943) and Alderfer (1969) among others. That is, the importance of core needs ensuring social reproduction, which, in turn, is accounted for as the primary motive for human action. Respective core needs are physical well-being referring to the satisfaction of basic needs, and social appreciation representing the desire of personal growth. However, neither need is easily satisfied nor directly accessible, but rather require the investment of other resources: for example, time.

To address the question of according to which rationale humans act on a motivation, EV- Theory was introduced. At the core of the theory is the idea that an individual would choose the best cost-benefit scenario from a set of possible action alternatives, considering the least amount of risk. Applied to the field of education, the work of Esser (1999) is pioneering, given that he incorporated meaningful approaches of others such as Boudon (1974), Erikson & Jonsson (1996), and Breen & Goldthorpe (1997), and introduced an EV-Theory based framework that considers several essential parameters guiding individual educational decision making. Parameters are the motivation for educational investments as well as associated risks. Furthermore, Esser (1999) emphasized socioe-

conomic background as an essential factor, and argued that families with a lower social background are more likely to refrain from educational investments, given that the utility of educational success is lower compared to the required costs.

Referring back to the Rubicon Model as an overarching framework, EV-Theory is included in the model, aiming to substantiate the first phase, that is the pre-actional phase of motivation. This is meaningful for two reasons. First, the Rubicon Model itself is based on EV-Theory guided concepts presented by Atkinson (1957, 1964) among others, and second, EV-Theory proved to be a widely accepted rationale for human actional decision making. That applies to both general choices of action and educational choices, respectively. Both are beneficial to the scope of this study, that is investigating the educational choices of socioeconomic low-performers in South Africa. Figure 3 shows the adapted Rubicon Model (fig. 2), substantiated with Expectancy-Value Theory.

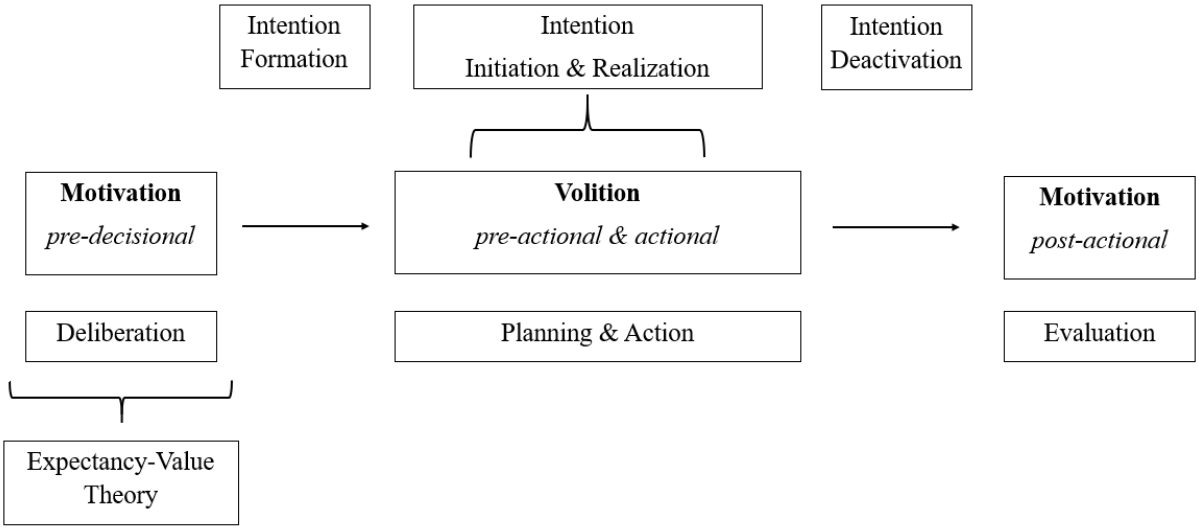

Figure 3: The adapted Rubicon Model of Action Phases (after J. Heckhausen & Heckhausen, 2008), substantiated with theories of rational decision making.

3.3. Parental Involvement

3.3.1. Theories of Parental Involvement

Along the lines of the Rubicon Model, the following section focuses on the association of the premises to models of parental involvement. Not only do pioneering models of parental involvement mirror the phases of the Rubicon Model, motives for parental involvement can also be seen in the light of rational decision making, particularly under the consideration of socioeconomic status.

Definition of Parental Involvement

To tie in with the aforementioned systematization of core needs motivating human action, parental involvement first and foremost includes the satisfaction of children's basic necessities such as food, clothing, shelter, health, and safety (Epstein, 1987, p. 121). However, apart from that, there is no universal understanding of parental involvement (W. Fan & Williams, 2010; X. Fan & Chen, 2001). For example, Maccoby & Martin (1983) describe parental involvement as the degree to which a parent is "committed to his or her role as a parent and to the fostering of optimal child development" (Maccoby & Martin, 1983, p. 48). Grolnick & Slowiaczek (1994) apply a broader yet economically focused definition, describing parental involvement as the "dedication of resources by the parent to the child within a given domain" (Grolnick & Slowiaczek, 1994, p. 238). Wilder (2014) refers to a broader definition of the construct provided by Jeynes (2007), framing parental involvement as the participation of parents "in the educational processes and experiences of their children" (Wilder, 2014, p. 390). Finally, attempting to summarize different definitions of parental involvement, Fan & Chen (2001) state that "despite its intuitive meaning, the operational use of *parental involvement* has not been clear and consistent" (X. Fan & Chen, 2001, p. 3); and thus, come to the conclusion that parental involvement "subsumes a wide variety of parental behavioral patterns and parenting practices", pointing towards the multidimensional nature of the construct (X. Fan & Chen, 2001, p. 3). On that note, parental involvement can be defined as a "catch-all term for many different activities including 'at home' good parenting, helping with homework, talking to teachers, attending

school functions, through to taking part in school governance" (Desforges & Abouchaar, 2003).

Resulting from the disagreement on a definition of parental involvement, researchers are also heterogeneous regarding the question of which activities educational parental involvement entails, yielding multiple definitions and conceptualizations of the construct, and frequently causing inconsistent empirical findings related therewith (Castro et al., 2015; X. Fan & Chen, 2001; Wilder, 2014).

For example, Grolnick & Slowiaczek (1994) suggest a threefold approach to parental involvement, distinguishing between behavioral, personal and intellectual aspects (Grolnick & Slowiaczek, 1994). A more detailed approach has been presented by Epstein (1987, 1992), who suggested four dimensions, that are basic obligations of parents, school-to-home communications, parental involvement at the school, and parental involvement in learning activities at home (Epstein, 1987, 1992).[27] Meta-analyses on parental educational involvement identified two overarching concepts to be most relevant, which are home-based and school-based parental involvement activities, empirically measured from the parental perspective (Punter et al., 2016).

Aimed at exploring parental involvement activities, Fan & Chen (2001) conducted a meta-analysis on 25 empirically relevant studies regarding parental involvement in association to student academic achievement, and found several dimensions of parental involvement activities, which were dominantly applied. Among those are actions related to parent-child communication, home supervision, and school contact and participation (X. Fan & Chen, 2001). Activities that typically pertain to home-based parental involvement are helping with homework, (early) literacy activities, discussing student's schooling, parenting style, parental monitoring and rule-setting, and ensuring school readiness. School-based

[27] In 1995, Epstein further distinguished six pillars of parental involvement, that are parenting, communicating, learning at home, volunteering, decision-making and community connections. Adding to respective concepts, Fan & Chen (2001) suggested a seven-stage-concept of parental involvement. However, a further distinction of dimensions is not meaningful for this study and is therefore not consolidated here (W. Fan & Williams, 2010).

parental involvement commonly includes the attendance of parent meetings, participating in school activities, as well as parent-teacher communication (Punter et al., 2016, p. 10). Based on the work of Epstein & Salinas (1992), and Shumow & Miller (2001), Walker et al. (2005) suggested a five-item home-based and a five-item school-based scale for parental involvement, specifically addressing the frequency according to which certain actions towards children's educational attainment are carried out by the parents. Regarding home-based involvement, Walker et al. (2005) asked parents how often they talk to the child about the school day, supervise the child's homework, help the child study for tests, practice spelling, math or other skills with the child, and read with the child. For school-based involvement, the authors were interested in how often parents help out at the child's school, attend special events at school, volunteers to go on class field trips, attend PTA meetings, and go to the school's open house.

Thus, Walker et al. (2005) among others assume a multidimensional construct of parental involvement. In contrast to this, some researchers argue for a unidimensional operationalization, given that certain dimensions of parental involvement are considered to be more influential than others (K. Singh et al., 1995), and raise the question of whether parental involvement would be better reflected through single items rather than multidimensional constructs (X. Fan & Chen, 2001). Hence, a practical approach to the discourse was provided by Fan & Chen (2001), and Fan & Williams (2010), who proposed acknowledging the multidimensional nature of parental involvement through measuring different dimensions of (home-based and school-based) parental involvement separately (W. Fan & Williams, 2010; X. Fan & Chen, 2001).

The Model of Parental Involvement Process

Although strongly supported with regard to operationalizing forms of parental involvement, criticism on persisting concepts as presented by Grolnick & Slowiaczek (1994) and Epstein (1987, 1992), among others, were raised regarding the lack of explanations as to why parents get involved as well as how their involvement affects student outcomes in learning (Walker et al., 2005). Aimed at addressing the issue, Hoover-Dempsey & San-

dler (1995, 1997) introduced a comprehensive framework, namely the *Model of Parental Involvement Process* (Walker et al., 2005). The model consists of a five-level framework aiming to "more fully understand families' contributions to student outcome" (Walker et al., 2005, p. 85). Figure 4 shows the original Model of Parental Involvement Process.

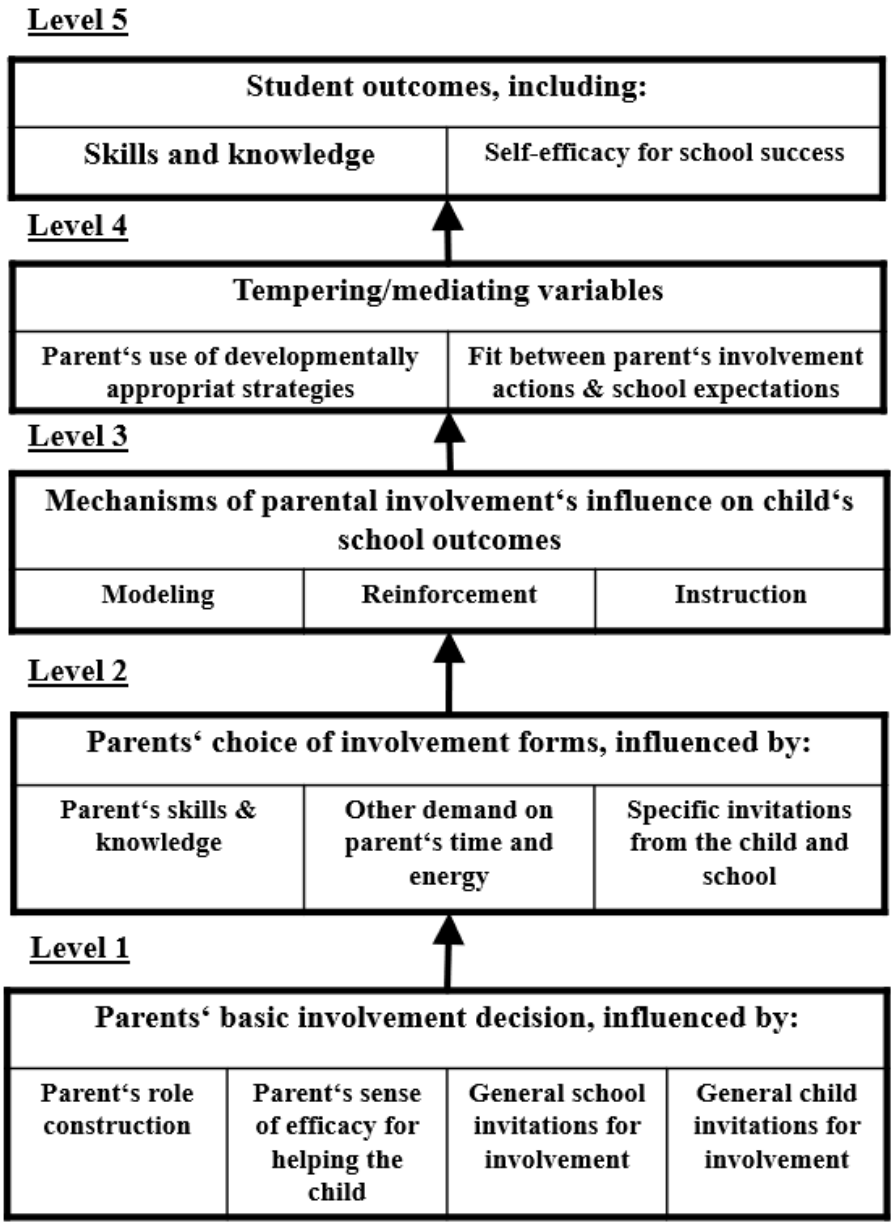

Figure 4: Hoover-Dempsey & Sandler's (1995, 1997) original Model of Parental Involvement Process (cited after Walker et al., 2005).

The model consists of five levels reflecting the process of parental involvement.
The first level is subdivided into four psychological factors influencing parent's basic involvement decisions. Factors include parents' role construction, parents' sense of efficacy for helping the child, as well as general school invitations and general child invitations for involvement. The second level focuses on factors influencing parents' choice of involvement forms, given that the decision to get involved was already taken on level one. Factors include parents' skills and knowledge, demands on parents' time and energy, and specific invitations from the child and school. The third level represents mechanisms of parental involvement and its influence on the child's school outcomes. The mechanisms are modeling, reinforcement and instruction. The fourth level summarizes factors that are assumed to mediate parental involvement, for example, parent's use of developmentally appropriate strategies, and the fit between parents' involvement actions and school expectations. The fifth and final level of the model points towards the effects of parental involvement reflected by students' outcomes, including skills and knowledge as well as self-efficacy for school success (Walker et al., 2005).

However, the Model of Parental Involvement Process shows some deficiencies. In a retrospective, Walker et al. (2005) conclude that the model was based on "current 'best guesses' " (Walker et al., 2005, pp. 85–86) and provided a "theoretical 'map' that connected existing bodies of knowledge" at the time (Walker et al., 2005, p. 87). Furthermore, the model lacks a clear approach to operationalizing and empirically verifying its major elements (X. Fan & Chen, 2001). Walker et al. (2005) aimed at addressing this issue, and started out with providing a list of items that were used to test the first two stages of the model. Items were retrieved from the existing body of research as well as developed for the purpose of their study based on survey data collected from parents of students in U.S. elementary schools. Finally, Walker et al. came to the conclusion that certain categories of the original models' first two levels needed to be collapsed, leading to a condensed version thereof. Figure 5 shows the revised Model of Parental Involvement Process.
Starting from the bottom Walker et al. (2005) defined three latent constructs at level one,

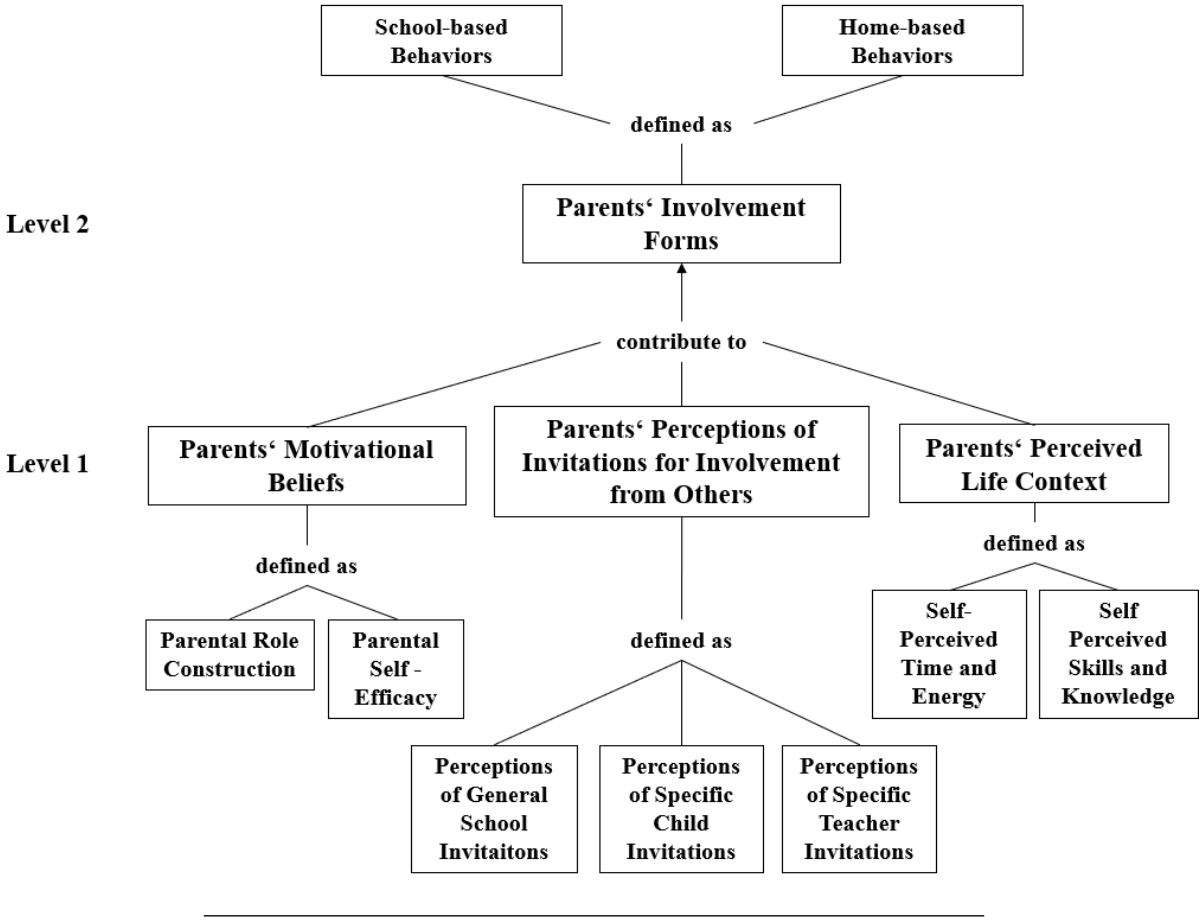

Figure 5: The revised Model of Parental Involvement Process (after Walker et al., 2005).

namely parents' motivational beliefs, parents' perceptions of invitations for involvement from others, and parents' perceived life context. The three constructs summarize level one and two of the original Model of Parental Involvement Process after Hoover-Dempsey & Sandler (1995, 1997). Parents' motivational beliefs summarize parental role construction and parental self-efficacy; both are influenced by a general understanding of what parents believe they should and can do in terms of supporting educational success of their children. Parents' perception of invitations for involvement from others refers to the perception that individual involvement is "sought, welcomed, and valued by the child, the child's teacher, and the child's school" (Walker et al., 2005, p. 93). Lastly, parents' perceived life context refers to parental involvement being influenced by the self-perceived availability of time and energy as well as skills and knowledge of parents. According to Walker et al. (2005), "these three overarching constructs represent the psychological underpinnings of parents' involvement behavior" (Walker et al., 2005, p. 87). Furthermore, the authors defined the constructs at level one to be related to parental motivation and resources predictive of parental involvement forms. This is in line with EV-Theory, stating that a decision to invest in something is always influenced by the availability of individual resources.

At level two, Walker et al. (2005) collapsed the independent measures of the original model, that are measures predicting parents' basic involvement decisions at level one, and factors predicting parent's choice of involvement forms at level two, leading them to downsize the model and introduce parents' involvement forms, which are further subdivided into school-based and home-based involvement activities (Walker et al., 2005). Walker et al. (2005) thereby picked up on the distinction originally suggested by Epstein (1987, 1992).

In their model, Walker et al. (2005) deliberately did not go beyond the investigation of why and how parents get involved. However, as incorporated in the original Model of Parental Involvement Process after Hoover-Dempsey & Sandler (1995, 1997), the association of parental involvement forms with student outcomes in learning cannot be untended for a comprehensive display of parental involvement.

3.3.2. Additional Concepts associated with Parental Involvement

Although the (revised) Model of Parental Involvement Process by Hoover-Dempsey & Sandler (1995, 1997), and Walker et al. (2005) is considered a comprehensive approach to parental involvement, there are some concepts that are considered to be associated with it.

For instance, aiming to operationalize motives for parental involvement at level one, aspects were included to the Model of Parental Involvement Process that deliberately predict parental involvement (Hoover–Dempsey et al., 2005; Walker et al., 2005). However, the current body of research points towards the inclusion of additional aspects, one of which are parental educational aspirations of the child. For example, Fan & Chen (2001) found in their meta-analysis sufficient empirical evidence that parental educational aspirations are a relevant predictor for parental involvement and student outcomes. Furthermore, results showed that among all aspects considered with regard to parental involvement, aspirations for the child's education showed the strongest association with student outcomes (X. Fan & Chen, 2001). For comparison, a similarly strong result was found by Peng & Wright (1994), who analyzed the association of parental educational aspirations with student performance across ethnic groups in the U.S. based on the *National Education Longitudinal Study* of 1988 (NELS:88). Results showed that correlations between parental educational aspirations and student outcomes were significantly positive within and across groups of interest (Peng & Wright, 1994). Thus, parental educational aspirations are generally considered a relevant motive for parental involvement and student performance, respectively.

Respective empirically substantiated assumptions are also supported in theoretical concepts such as EV-Theory where preservation of social status is the essential motive for decisions on educational investments. Thus, parental educational aspirations are considered a relevant factor in the theoretical framework of this study. On that note, it is assumed that the higher the need for status preservation is, the higher the parental educational aspirations for student achievement are. According to EV-Theory, this is particularly true for parents of higher-socioeconomic status.

Furthermore, as the original Model of Parental Involvement Process stipulates, parental involvement activities affect not only students' skills and knowledge, but also their sense of self-efficacy in learning. The latter is "a performance-based measure of perceived capability" (B. J. Zimmerman, 2000, p. 82). It originates from theoretical approaches of *Social Cognitive Theory* and *Social Learning Theory* (Bandura, 1997), and refers to "individuals' beliefs in their ability to produce desired results (...) as well as to learn and perform" (W. Fan & Williams, 2010, p. 56). Furthermore, researchers found that self-efficacy is positively associated with student outcomes in learning, given that those who perceive themselves as self-efficacious are more resilient to and proactive in challenging learning situations (W. Fan & Williams, 2010). Taken together, student self-efficacy has consistently been shown to be a strong predictor for student outcomes in learning (Bong, 2008; B. J. Zimmerman, 2000), and parents can foster students' sense of self-efficacy by becoming more involved (Eccles et al., 1998; Grolnick & Slowiaczek, 1994).

High student self-efficacy is also associated with high intrinsic motivation to learn, and taken together, both of these personal motivational beliefs are predictive for student performance (Bong, 2008). Aiming for a definition of intrinsic motivation, Fan & Williams (2001) argue that those "who demonstrate intrinsic motivation engage in academic tasks due to the enjoyment of the tasks and the desire to learn" (W. Fan & Williams, 2010, p. 57). Hence, considering *Cognitive Evaluation Theory* after Deci & Ryan (1980, 1985), Fan & Williams (2001) make the case for extending the scope of parental involvement regarding its effect on students' intrinsic motivation to learn as well (Bong, 2008; W. Fan & Williams, 2010).

However, on that note, it must also be noted that parental involvement is known to have controlling effects on student performance. Whereas some parental practices, such as positive reactions to grades, are known to potentially foster higher intrinsic motivation among students, other activities, for example getting involved with homework, may cause the opposite (W. Fan & Williams, 2010).

3.3.3. Implications for the Study

To summarize, the (revised) Model of Parental Involvement Process developed by Hoover-Dempsey & Sandler (1995, 1997) and revised by Walker et al. (2005) can be viewed in the context of the Rubicon Model of Action Phases. That is, the pre-decisional phase is equivalent to parents' motivation to get involved, the volitional phase aligns with parents' involvement forms, and the post-actional phase is represented by the evaluation of parental involvement through its effect on student outcomes in learning.

Factors of pre-actional motivation are parents' motivational beliefs, parents' perceptions of invitations for involvement from others, and parents' perceived life context. According to the model, respective aspects are predictive of parents' involvement forms, and hence, their volitional educational investment behavior. Furthermore, parents' motivational beliefs and parents' involvement forms are associated with the availability of different resources, such as skills and knowledge, and time (Walker et al., 2005). Respective aspects are considered predictive of educational investments within the framework of EV-Theory as well. Moreover, it is assumed that parents of higher socioeconomic status are better equipped with respective resources, and have more reason to invest them. The latter is due to the elevated need for status-preservation.

Parents' involvement forms defined as school-based and home-based activities represent the volitional phase of the Rubicon Model. This is because parents' involvement forms are often reflected by the frequency of parental involvement activities. Recall, home-based parental involvement summarizes actions such as talking to the child about the school day, and supervising the child's homework.

Lastly, the evaluation of volitional behavior, that is post-actional motivation, is represented through student outcomes in learning, defined as skills & knowledge, students' self-efficacy, as well as the additional concept of intrinsic motivation, acknowledging the association with students' self-efficacy.

Justified by the aforementioned interconnection of both frameworks, the (revised) Model of Parental Involvement Process by Hoover-Dempsey & Sandler (1995, 1997) and Walker

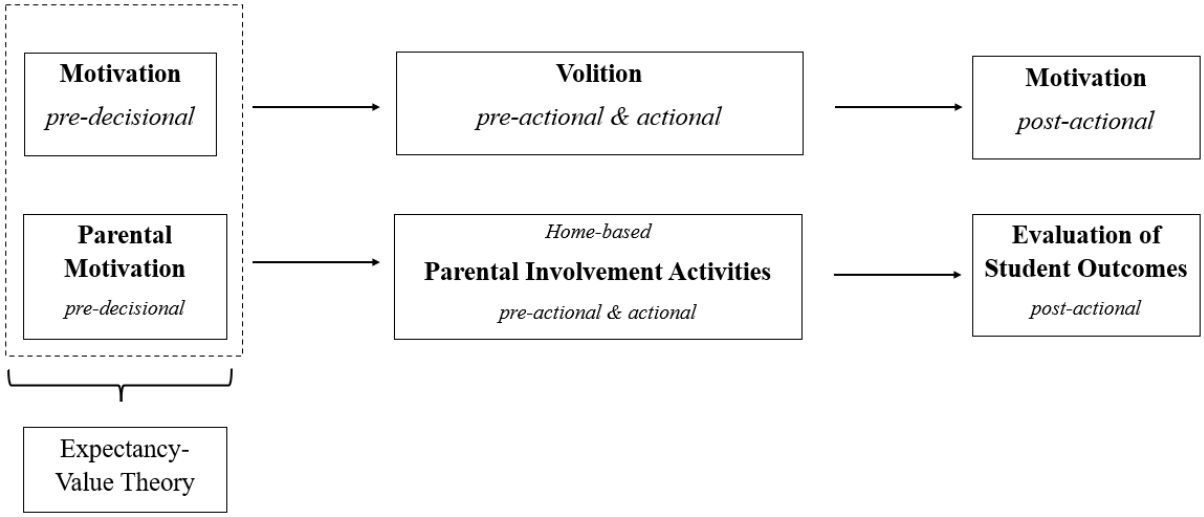

Figure 6: The theoretical framework of this study, based on the adapted Rubicon Model of Action Phases (after J. Heckhausen & Heckhausen, 2008) and the (revised) Model of Parental Involvement Process (after Walker et al., 2005), substantiated with theories of rational decision making. Own illustration.

et al. (2005) is mirrored onto the Rubicon Model. Figure 6 shows the adapted Rubicon Model, applying the levels of parental involvement as summarized in this section.

In order to operationalize the pre-decisional and volitional phase of parental involvement forms, this study focuses certain aspects of the (revised) Model of Parental Involvement Process. This is due to two reasons. First, the data set of prePIRLS 2011, which is utilized for this study, limits the operationalization of the theoretical model, and thus, focuses invitations for involvement from others at the pre-decisional phase. Second, the scope of this study is on self-perceived educational investment behavior and involvement activities, respectively. Consequently, forms of home-based involvement are chosen over forms of school-based parental involvement to represent volitional behavior. On that note, this study further distinguishes between different home-based involvement activities (sec. 8.2.3), and hence, the terminology is changed, referring to parental involvement activities rather than forms. Respective limitations are considered in figure 6 as well.

3.4. Summary

This study aims at analyzing motivation, volition and the evaluation thereof in the scope of (parental) educational investment behavior among socioeconomic low-performers in South Africa. To that end, three overall concepts as well as the associations thereof have been presented.

First, the core framework of this study, the Rubicon Model of Action Phases introduced by Heckhausen & Gollwitzer (1987) was introduced as the foundation for the theoretical framework of the study. The model incorporates three major phases of human action behavior, which are pre-decisional motivation, pre-actional and actional volition, and the post-actional evaluation thereof. The phases of the model are understood as a causal chain of events. As pointed out, the Rubicon Model is applicable to various fields of research. To fit the purpose of this study as well as the available sources of data, the model was slightly adapted in the sense that the pre-actional and actional phases were collapsed and subsumed as volitional behavior, accordingly.

Second, core needs of human action as well as the premises of their saturation have been briefly addressed. To tie in with the question according to which paradigms humans chose to act or not act on a certain need, Expectancy-Value Theory has been introduced. Pointing to the role of EV-Theory in educational decision making, and how respective mechanisms of cost-benefit calculations are decisive to educational attainment, the importance of socioeconomic status has been reflected on as well. In this study, EV-Theory is used to underpin the first phase of the Rubicon Model, that is the pre-decisional phase, aimed at allowing for theoretically sound assumptions about the motivation for (parental) educational investment behavior.

Third, theories of parental involvement introduced by Hoover-Dempsey & Sandler (1995, 1997) and revised by Walker et al. (2005) have been consulted. Whereas the original Model of Parental Involvement Process aimed at comprehensively depicting the mechanisms of parental involvement, Walker et al. (2005) specifically emphasized the aspects of parental motivation for involvement and provided empirical evidence for their validity.

Though, to account for limitations resulting from prePIRLS and the scope of this study, selected motivational aspects of parental involvement as well as home-based involvement activities are focused. Hence, the following sections particularly report on these.
Lastly, the relevance of additional concepts associated with parental involvement have been introduced. This concerns parental educational aspirations functioning as a motive for parental involvement at the pre-decisional phase; and students' intrinsic motivation as a complementary concept to students' personal motivational beliefs, such as self-efficacy, according to theories of Social Cognitive Theory and Social Learning Theory after Bandura (1997), as well as Cognitive Evaluation Theory after Deci & Ryan (1980, 1985).

4. State of Research

The following section summarizes findings of empirical research on the topic of educational investment behavior and parental involvement as a form thereof. Although, educational investments are often taken by officials and governmental bodies, the focus of this study calls for emphasis on individual rather than institutionalized educational investments. On that note, the following section focuses individual educational investment behavior, as well as its association with socioeconomic status, and its relevance for the persistence of social inequality. That is, with special attention to the South African context, and beyond.

The presentation of the state of research is oriented along the lines of this study's theoretical framework, outlined in section 3. Hence, only certain aspects of the motivational and volitional phase of parental involvement are emphasized. Furthermore, the state of research is presented thematically rather than chronologically, given that the interest lays with the content rather than the development of the subject. To that end, findings on educational investments as a driver for educational inequality are summarized first, followed by the state of research on parental involvement as a form of educational investment. Here, motivational factors and volitional aspects of parental involvement are focused, and parental involvement is viewed in the light of socioeconomic status and different cultural contexts. Lastly, this section is concluded by positioning this study within the current state of research, and research gaps are identified.

4.1. Educational Investments

4.1.1. Definition of Educational Investments

> "*Differences of 'education' today (...) are undoubtedly the most important actually state-forming difference. [Unterschiede der ‚Bildung' sind heute (...) zweifellos der wichtigste eigentlich ständebildende Unterschied].*" (Max Weber (1922), cited after R. Becker & Lauterbach, 2016, p. 3).

Although the statement was made a century ago, its message still accurately describes the current effect of education and educational opportunities on modern societies. That

is, education is not only considered essential for the accessibility of life opportunities, but also the accumulation of economic, cultural and social capital and the determination of socioeconomic status (R. Becker & Lauterbach, 2016; Maaz et al., 2008; Stocké, 2010). For example, individuals with higher educational degrees tend to have higher income (Boockmann & Steiner (2006), Auld & Sidhu (2005), cited after Stocké, 2010), lower risk for unemployment (Kettunen (1997), cited after Stocké, 2010), and better overall health (Sander (1998), cited after Stocké, 2010). Hence, educational investments are prone to having a remedial effect on educational disparities and states of social inequality.

Based on theories of rational educational decision making, the concept of educational investments has been studied extensively, especially with regard to its effect on mitigating educational and societal disparities. Along the lines of Gary Becker's *Economic Theory of Human Capital* (1975), educational investments typically refer to the allocation of monetary resources towards education, that is to defray opportunity costs as well as direct costs of schooling (Catsiapis, 1987; Stocké, 2010). Opportunity costs on the other hand result from labor market absence and the associated loss of income while attaining to education (Esser, 1999; Stocké, 2010). Direct educational costs are measured through the financial load that must be accommodated when taking an educational investment, for example attaining a degree in higher education. Direct costs entail school fees and expenses for learning materials among others. Furthermore, empirical studies pointed toward the importance of considering the availability of excess money at the end of the month, and number of children in the family for the evaluation of direct costs (R. Becker, 2003). With regard to the latter, a study by Peraita & Pastor (2000) provided evidence for the assumption that resources per capita potentially available for educational investments are lower in families with more children (Peraita & Pastor, 2000).
In addition to monetary resources required for educational investments, researchers also emphasized the factor of time (G. S. Becker, 1962; Esser, 1999; J. Heckhausen & Heckhausen, 2008; Nerdinger, 1995). For example, time is relevant to the decision of a particular educational pathway such as a higher school track. That is, the attainment of a

higher educational degree requires more time than the attainment of a lower or even no educational degree. Thus, the longer an educational pathway takes, the more time one is obliged to invest for it. With that in mind, it is unsurprising that families' socioeconomic status is positively associated with the duration of students remaining in education (Conley, 2001); and thus, with the probability of pursuing a degree in higher education (Light & Strayer, 2000). To that end, the proverb 'time is money' holds true.

Referring back to the initial quote by Max Weber (1922), educational investments are generally considered to enhance one's life opportunities, and contribute to the accumulation of economic, cultural and social capital, the enhancement of socioeconomic status, accessibility of life opportunities, and the mitigation of social inequalities (R. Becker, 2000). However, as stated in the introduction of this study and along the lines of rational educational decision making (sec. 3.2.2), socioeconomic status is not only determined by educational investments, but also functions as a predictor thereof (Stocké, 2010). Thus, it affects rational educational decision making and, in fact, enhances the probability of manifesting educational inequalities rather than mitigating them (R. Becker & Lauterbach, 2016).

4.1.2. Educational Investments as a Driver for Inequality

International Findings

The effect of educational investments on behalf of status preservation and as a function of socioeconomic status have been highlighted by researchers around the globe, especially with regard to educational systems in the U.S. and Europe, and at points of transition in the school system.

For instance, Becker (2000) found that, in Germany, where the educational system is stratified, higher socioeconomic status and the goal to preserve thereof is likely to enhance the probability of students continuing on a higher school track after primary education (R. Becker, 2000). Accounting for different time points, results furthermore confirmed

that the elevation of socioeconomic status over time entails a shift in factors determining educational decision making. For example, increasing economic prosperity leads to educational costs being less decisive for educational investments, and status preservation becomes a more important argument (R. Becker, 2003). Hence, studies by Becker (2000, 2003) empirically validated theoretical assumptions of rational educational decision making. Referencing respective findings among others, Becker (2016) pointed towards the importance of track choices at points of transition for the establishment and preservation of educational disparities. This is because educational decisions of parents and students, especially at the end of primary school and at points of transition, proved to be decisive for educational careers.

On that note, Stocké (2007) utilized data from Germany, analyzing whether the individual assessment of costs, utilities and chances for success in terms of obtaining a certain educational degree is subjected to socioeconomic status. Results showed that individuals of higher socioeconomic status perceive the attainment of the German university entrance qualification to be less costly yet highly beneficial as well as likely to be accomplished. Furthermore, group differences in the perception of expected cots have been linked back to differences in income and number of children in the family (Stocké, 2007a). A related study by Stocké (2007) on parents of students in German primary schools showed that parental assessments of the probability that certain educational degrees will ensure status preservation for their children increase with the type of degree, and decrease with family socioeconomic status (Stocké, 2007b). Hence, status preservation is tied more strongly to educational attainment among socioeconomically high-performing families. This is due to the fact that ensuring a certain standard of living requires more resources, which are more likely to be accumulated if a certain educational attainment is ensured. Thus, the assumption of rational educational decision making being associated with the motive of status preservation and socioeconomic status holds true (Stocké, 2010).

Although treated as the main determinant of rational educational decision making, studies consciously added to the motive of fear of loss of status by shedding light on the effect of values and beliefs in association with socioeconomic status. This approach contradicts the general premises of rational decision making, assuming that ability and resources vary by socioeconomic status, but not so educational values and beliefs (Breen & Goldthorpe, 1997; Sullivan, 2006).

On that note, Sullivan (2006) emphasized the effect of self-perceived student ability for the repetitiveness of educational disparities. It was shown that, on average, graduating students in the U.K. subjectively overestimate their ability to achieve. Interestingly, effects were stronger among students from socioeconomically high-performing families (Sullivan, 2006). Assuming that self-perception of educational attainment is relevant to the perception of the probability for educational success and, hence, fosters decisions for the attainment of higher educational degrees, it can be concluded that respective effects add to the continuity of gaps in educational attainment and life opportunities (Stocké, 2010).

Findings on the South African Context

As shown in section 2, unequal life opportunities in South Africa are reflected through various identifiers. These are ethnicity, native language, area of residency, and educational opportunities among others. With regard to the latter, several economically oriented studies on the development and manifestation of educational inequalities in South African have been published. Again, the availability of capital functioning as a decisive factor for the persistence of unequal life opportunities is at their core.

To narrow the racial gap in South Africa, Branson et al. (2012) identified the attainment of post-secondary education as an important factor for the accumulation of capital and life opportunities. As shown in section 2, differences in educational attainment are evident between groups of Whites and Black-Africans. With Whites being enabled to more total years of schooling, especially in post-secondary education, Branson et al. (2012)

found that they are less prone to income losses and unemployment. Vice versa, Black-Africans improved in primary education over time, but stayed behind their counterparts in completing post-secondary education. Thus, they are more often affected by unemployment and poverty (Branson, Garlick, et al., 2012). Observing years of schooling in association with income development, the authors unsurprisingly found a co-variation of the two. Furthermore, Black-Africans with 15 years of schooling, that is a completed university degree, showed similar educational returns in earnings as Whites with only 12 years of schooling completed (Branson, Garlick, et al., 2012). That is, the historically and politically disadvantaged group of Black-Africans must allocate more time and monetary resources towards educational attainment in order to achieve the same income level as Whites in South Africa. Finally, this chain of effects adds to the continued persistence of racial gaps and unequal life opportunities.

In addition to the aforementioned are school dropouts as well as grade repetition or progression. They are considered to have decisive effects on the reproduction of educational and social inequality. This is because the prolonging or even discontinuation of formal education leads to disparities in levels of schooling and educational attainment, and fosters the persistence of unequal life opportunities (Lam et al., 2011). In 2000, class repetition rates in South Africa accumulated to an average of 9% between grades 1 - 7. Against this background, Ardington et al. (2011) and Lam et al. (2011) analyzed longitudinal data from the *Cape Area Panel Study* (CAPS) and found with regard to racial differences in educational advancement throughout primary and secondary schooling in South Africa that Black-Africans were lagging behind their counterparts substantially. For example, whereas 82% of Whites progressed by three grades in three years (i.e., from 2002 and 2005), only 27% of Black-African students achieved the same. To that end, Black-African students showed high rates of grade repetition as well as school enrollment well above the average student age at the end of post-secondary schooling, that is when taking the NSC or Matriculation respectively (sec. 2). Aimed at explaining differences in school advancement, Ardington et al. (2011) and Lam et al. (2011) considered the effect of

socioeconomic background factors. They found that family characteristics, which are household income and parental educational status, educational performance as well as past educational pathways account for the racial gaps in grade progression. Including school level characteristics, which are type of school, student-teacher ratio, and school fees as a proxy for school quality, it was also found that those do not substantially weaken the effect of individual student characteristics on grade progression. Hence, racial gaps in school progression are rather directed by individual background factors than school quality (Ardington et al., 2011; Lam et al., 2011).
In summary, results shown by Ardington et al. (2011) and Lam et al. (2011) hint towards the persistence of socioeconomic status determining educational decision making in South Africa. That is, students' cultural and social capital is decisive to grade progression and successful completion of secondary schooling.

To conclude, educational investments have certainly been found to be beneficial for the purpose of mitigating social inequalities. However, the motivation and capability for taking respective investments depend on individual resources, namely time and monetary means, as well as socioeconomic status. As shown in studies by Lam et al. (2011) and Ardington et al. (2011), mechanisms of rational educational decision making proved to be fitting for the South African context as well.

4.2. Parental Involvement as a Form of Educational Investment

> "*Parents who invest less time and energy in raising their children, or in guiding their offsprings' educational progress, curtail their children's human capital and future earning power, causing the reproduction of inequality across generations.*" (Mulligan (1997), cited after Attewell & Newman, 2010, p. 8)

On that note, parental investments in education are investments in children's cultural and social capital. The following section emphasizes parental involvement as a form of educational investment behavior.
Bridging the gap between rational decision making and parental involvement, the expectation of an outcome in education to be successful is strongly associated with the motivation

and capabilities of parents to support an educational decision (Stocké, 2010). Along those lines, the following section provides an overview of the state of research on home-based parental involvement. Yet, to tackle the vast amplitude of research on the subject (e.g., Castro et al., 2015; X. Fan & Chen, 2001), and in alignment with the theoretical framework of this study (sec. 3), emphasis is put on motivational and volitional aspects of educational involvement, namely parental perceptions of invitations for involvement from others and parental educational aspirations, as well as home-based parental involvement activities (sec. 3.3).

Taken together, studies on home-based parental involvement from the perspective of parents and in association with different aspects of student outcomes in learning, while considering effects of socioeconomic status, are most conducive to positioning this work within the body of research. To that end, a recently published study by Punter et al. (2016) summarizing relevant work on parental involvement with regard to its association with student outcomes in reading has been chosen as a point of reference for the literature review (Punter et al., 2016). The authors included research that focused on primary school students, given that effects of parental involvement are more relevant to the achievement of younger students (Domina, 2005; X. Fan & Chen, 2001; Jeynes, 2005; K. Singh et al., 1995), as well as studies that considered the effect of socioeconomic status and addressed additional outcome variables, for example student motivation. At the same time, studies on parental involvement in secondary education (e.g., Hill & Tyson, 2009), from student or school perspective (e.g., Domina, 2005; Dumont et al., 2012; Mattingly et al., 2002; Sénéchal & Young, 2008), or with regard to other dimensions of parental involvement (e.g., Castro et al., 2015; Domina, 2005; Dumont et al., 2012; Hill & Tyson, 2009; Okpala et al., 2001; Wilder, 2014) are largely disregarded here.

Finally, the body of research roughly comprises studies of the last 25 years, given that the topic of parental involvement was heavily researched during that time span. More recent work is included if available. Due to its vast extent, the literature review does not claim to be exhaustive but merely comprehensive in the sense that the most relevant findings

are summarized. That being said, studies built on the (revised) Model of Parental Involvement Process are preferentially included, as they relate to the operationalization of parental involvement as utilized for this study.

4.2.1. Motivational Factors for Parental Involvement

Although parental involvement activities are considered important for effective schooling, little is known about the motives for parents to get involved (Green et al., 2007). On that note, Walker et al. (2005) presented the empirically revised Model of Parental Involvement Process, yielding in three motivational dimensions predictive of parental involvement activities. Those are parents' motivational beliefs, parents' perceived life context, and parents' perceptions of invitations for involvement from others (Walker et al., 2005). Only the latter are emphasized in this study (sec. 3.3; sec. 8.2.3).

Perceptions for Involvement from Others

Following a long tradition on parents' perceptions of invitations for involvement (e.g., Epstein, 1986; Griffith, 1998; Hoover–Dempsey & Sandler, 1997), they are considered "key motivators of parents' decision to become involved" (Hoover–Dempsey et al., 2005, p. 110). The positive effect is premised on the associated sense of parental involvement activities being perceived as welcomed, valued, important and expected (Hoover–Dempsey et al., 2005). Hence, parents, who experience a school climate that fosters perceptions of a welcoming and valued environment tend to get involved in their children's education more frequently. This is especially advantageous for socioeconomically disadvantaged families in the primary educational sector, given that parents' volitional behavior can potentially be guided by schools' efforts to create a well perceived school climate and sense of appreciation (Griffith, 1998).

Several studies focusing the empirical validity of the psychological constructs of motivation for parental involvement along the lines of the revised Model of Parental Involvement Process were identified. Studies applying earlier versions of the dimension are excluded

here (Anderson, 2005; Campo, 2011; Desland & Betrand, 2005).

For example, Green et al. (2007) analyzed the effect of psychological constructs for parental motivation to become involved in their children's education while controlling for socioeconomic status and student age. The authors found that specific invitations for involvement from students and teachers are significantly associated with home- and school-based involvement. That is, student invitations significantly enhanced home-based involvement, and teacher invitations showed significant effects on school-based involvement. The authors also pointed towards the insignificance of general school invitations for involvement for home-based and school-based involvement activities (Green et al., 2007). Similar results were obtained by Walker et al. (2011), who conducted an exploratory study on parents' motivation for involvement in their children's schooling from grades 1 - 6 among 147 Latin-American families in the U.S., and found that home-based parental involvement was associated with specific invitations from students, whereas school-based involvement was associated with specific invitations from teachers. Again, general invitations for involvement by the school were not predictive of involvement activities (Walker et al., 2011).

On the other hand, results obtained by Reiniger & López (2017) confirmed the significantly positive effect of children's invitations for involvement, that is on home-based as well as school-based involvement activities (Reiniger & López, 2017).[28] However, results of multiple linear regression showed that student invitations are slightly more predictive for school-based involvement activities (.377) than for home-based involvement activities (.311). The authors also found that general invitations for involvement from the school are rather insignificant moderators compared to students' and teachers' invitations for involvement (Reiniger & López, 2017).

Taken together, studies presented here came to somewhat cohesive results regarding the

[28] Reiniger & López (2017) did not explicitly list aspects included to the dimensions of parents' perceptions for involvement from others. Given that the study is based on the revised Model of Parental Involvement Process, results are included here nonetheless.

effect of parental perceptions of invitations for involvement from others. That is, invitations from the child are overall more strongly associated with home-based involvement activities, and teachers' invitations proved to be more strongly associated with school-based involvement activities. Furthermore, invitations for involvement, regardless of who expresses thereof, showed to have positive effects on parental involvement, both at home and at school. This finding is reinforced by results presented by Desland & Bertrand (2005). Although focusing the effects of students' and teachers' invitations for involvement, and confirming the findings of aforementioned studies in that regard, they also found that any specific invitation for involvement enhances parents' educational involvement overall (Desland & Betrand, 2005). With regard to general invitations by the school, studies confirmed a negligible effect on parental involvement activities (e.g., Green et al., 2007; Reiniger & López, 2017; Walker et al., 2011).

Lastly, the body of research limits its scope to the first and second level of the (revised) Model of Parental Involvement Process. Hence, invitations for involvement from others are only analyzed with regard to their association with parental involvement. More holistic empirical approaches that extend the analyses to involvement activities being predictive of student outcomes in learning are not presented.

Parental Educational Aspirations

Not included in the Model of Parental Involvement Process yet widely discussed, especially with regard to student outcomes in learning, are parental educational aspirations (Bloom, 1980). Definitions on the construct are diverse, yet parental educational aspirations are typically operationalized through aspirations on the highest educational degree a child should obtain (e.g., Jodl et al., 2001; Peng & Wright, 1994; K. Singh et al., 1995).

Findings of widely referenced meta-studies on parental involvement indicate the relevance of parental educational aspirations as a concept associated with parental involvement (e.g., Castro et al., 2015; X. Fan & Chen, 2001; Jeynes, 2005, 2007; Wilder, 2014). However, research studies are indecisive on the role of parental educational aspirations.

For instance, earlier yet popular work on parental educational aspirations has been presented by Peng & Wright (1994) and Singh et al. (1995), who utilized data of the *National Education Longitudinal Study* of 1988 (NELS:88) conducted in the U.S., and aimed at analyzing student achievement from different angles. The NELS:88 is a representative study of 25,000 eighth-grade students as well as their parents, teachers and school administrators. Thus, the sample allowed for the analysis of diverse research questions based on achievement and background data. With regard to parental aspirations, Peng & Wright (1994) found that those are most strongly associated with overall student achievement in reading and mathematics (Peng & Wright, 1994). Similar results were obtained by Singh et al. (1995), who conducted a structural equation model and operationalized parental educational aspirations as a predictor of parental involvement as well as student achievement in different subjects. Results showed, that educational aspirations have a positive effect on student achievement as well as on parent-child communication about school. That is, even when controlling for socioeconomic background, ethnicity, and former levels of student achievement (K. Singh et al., 1995).

Furthermore, Seyfried & Chung (2002) analyzed longitudinal data from the *Seattle Social Development Project* (SSDP) with regard to the association of parental educational aspirations with student achievement among African American and European American eighth-grade students. Applying multiple regression analysis, they found that parental educational aspirations explained 36% of variance in students' GPA scores.[29] Results also showed that African American students received lower GPA scores, even though their parents showed similar levels of parental educational aspirations compared to their counterparts (Seyfried & Chung, 2002). A more recent study by Xu et al. (2010) analyzed the association of parental educational aspirations among others with student outcomes in reading and students' efficacy on self-regulated learning. Results showed that parental educational aspirations have a significantly positive effect on outcomes in learning (Xu et al., 2010).

[29] The Grade Point Average (GPA) is part of the U.S. academic grading systems.

Contradicting the positive notions of parental educational aspirations, a negative example was presented by Jodl et al. (2001). Based on the longitudinal study *Maryland Adolescent Development in Context* (MADIC), Jodl et al. (2001) analyzed the association of parents' aspirations with seventh-grade students' values and beliefs as well as their occupational aspirations. Applying linear and binary logistic regression, results showed that parental educational aspirations are not significantly associated with any of the considered outcome variables (Jodl et al., 2001).

Reviewing several studies as well as meta-analyses on parental educational aspirations, it became evident that the aspect has been considered inconsistently as a motivational factor for parental involvement. For instance, the study by Singh et al. (1995) made an effort to operationalize parental educational aspirations as a factor associated with parental involvement activities, and hence, acknowledge the differences between the two phases of motivation and action. On the contrary, studies by Peng & Wright (1994) and Xu et al. (2010) as well as the referenced meta-studies, did not consider a respective distinction between motivational and volitional aspects of parental involvement.

4.2.2. Volitional Aspects of Parental Involvement

> "*The idea that parental involvement has positive influence on students' academic achievement is so intuitively appealing that society in general, and educators in particular, have considered parental involvement an important ingredient for the remedy for many problems in education.*" (X. Fan & Chen, 2001, p. 1)

The importance of parental involvement as a component of effective education has been acknowledged by researchers, practitioners, and policy makers for a long time (Epstein, 1987, 1995, p. 119). Thus, it is unsurprising that the concept of parents who "can change their children's educational pathways by engaging with their children's schooling has inspired a generation of school reform policies." (Domina, 2005, p. 245). Prominent examples thereof are surely the widely referenced yet critically discussed policies implemented by the U.S. Department of Education, which are the *Every Student Succeeds Act* (ESSA) of 2015, and the preceding *No Child Left behind Act* (NCLB) of 2001, mandating schools

in the U.S. to enhance the involvement of parents in children's education (W. Fan & Williams, 2010).

As summarized in section 3.3, parental involvement is considered multidimensional, and lacks a concise definition. Researchers have been rather heterogeneous with regard to the empirical conceptualization of the construct, frequently causing inconsistent findings (e.g., Castro et al., 2015; X. Fan & Chen, 2001; Wilder, 2014). Along those lines, the following section focuses the association of multidimensional home-based parental involvement activities with student outcomes in learning. In certain cases, general parental involvement is reported if no additional information on the construct's operationalization are available. Also, findings that focus on single involvement activities, for example homework support (e.g., Hoover–Dempsey et al., 2001; Patall et al., 2008; Peng & Wright, 1994), are not addressed here.

Meta-Studies on Parental Involvement

With regard to parental involvement, several widely acknowledged meta-studies are available.

For instance, very frequently cited meta-analyses on parental involvement have been presented by Jeynes (2003, 2005, 2007, 2012). Focusing on primary education, the meta-analysis of 2005 comprises of 41 research studies, and showed that parental involvement is associated with student outcomes in learning in primary school. That is regardless of students' ethnicity and sex (Jeynes, 2005). Furthermore, results were reinforced in the meta-analysis of 52 studies in 2007, showing the same effect for students in secondary education (Jeynes, 2007).

Another popular meta-analysis has been presented by Fan & Chen (2001). They found that, although parental involvement has been heavily researched, there is a rather small number of empirical studies on the subject (X. Fan & Chen, 2001). Aiming to summarize those, Fan & Chen (2001) analyzed correlation coefficients of 25 research studies on parental involvement, and reinforced that, overall, parental involvement is indeed contributory to student achievement. This is despite the fact that individual studies showed

inconsistent findings in that regard. Furthermore, their analysis showed that effects of parental involvement are stronger with regard to overall student achievement, for instance the school GPA, compared to subject-specific learning outcomes, such as outcomes in mathematics or reading (X. Fan & Chen, 2001).

The referenced meta-analyses are positive examples, confirming the relevance of general parental involvement for overall student outcomes in learning. However, as Fan & Chen (2001) pointed out, single research studies sometimes come to different conclusions. Paying attention to that, the following summary acknowledges individual studies that find parental involvement activities to be positively as well as negatively associated with student achievement.

Individual Studies on Parental Involvement

In alignment with theoretical frameworks on parental involvement (sec. 3.3), Punter et al. (2016) retrieved a total of 22 studies, including meta-studies as well as individual research studies, and organized them by home-based and school-based parental involvement from the perspective of parents, students, and schools (Punter et al., 2016).[30] In summary, research studies on home-based parental involvement from the perspective of the parents showed to be in disagreement whether parental involvement is positively or negatively associated with students' performance in literacy (Punter et al., 2016).

Positive associations of global parental involvement and student achievement in reading were found by Bakker et al. (2007) and Kloosterman et al. (2011). Utilizing data from 218 parents collected in the Netherlands, Bakker et al. (2007) found that parents' perceptions of involvement at home were significantly predictive of students' performance in vocabulary tasks, mathematics and reading comprehension. Parents' perceptions of being informed by the school showed positive associations with performance in spelling tasks,

[30] Punter et al. (2016) stated that the goal of their pursuits was not to provide an exhaustive summary on the subject of parental involvement but to identify dimensions of parental involvement bearing the most significant potential to be influential on student achievement, and to analyze the effect of country-specific differences (Punter et al., 2016).

and mathematics (Bakker et al., 2007). Focusing on the Netherlands as well, Kloosterman et al. (2011) used panel data from the *Dutch Cohort Study Primary Education* (PRIMA), collected on a biannual schedule since 1996. The authors found that global parental involvement, consisting of information from parents and teachers on parental involvement in education, showed to be positively associated with student performance in reading in the first years of primary education (Kloosterman et al., 2011).

Studies that found a negative association of parental involvement with student outcomes in learning were shown by Barnard (2004), and Xu et al. (2010). For instance, Barnard (2004) analyzed data from the *Chicago Longitudinal Study* (CLS). Based on the cohort of 2000, results on 1,165 students at the age of 20 showed that parental involvement in school throughout primary education has a positive effect on students' remaining in school, finishing high school in time, and on the school-leaving degree (Barnard, 2004). Interestingly, results were only significant when using teachers' perceptions of parental involvement activities at school as predictors. On the contrary, parents' perception of home- and school-based involvement activities did not result in significant effects in any of the outcome variables. That is, although parents reported rather high rates of involvement at home as well as in school. Bernard (2004) named several possible explanations for the results, for example, little variance in parents' self-reported ratings of involvement activities. To address the issue, the author proposed to administer more items, at least on a five-point scale, to mirror parental involvement sufficiently (Barnard, 2004). Furthermore, Xu et al. (2010) analyzed data of fifth-grade students from the *Early Childhood Longitudinal Study* (ECLS-K) collected in the U.S. between 1998 - 1999. Despite considerable positive effects of parental involvement at school and in extracurricular activities, the authors found negative effects in the frequency of parental support in literacy and mathematics homework on student achievement in reading (Xu et al., 2010).

Studies on Parental Involvement based on PIRLS

After addressing results from international meta-studies and individual research studies, the state of research on parental involvement utilizing IEA's PIRLS data is presented last.

Attempting to assess the psychometric framework of (home-based) parental involvement, Punter et al. (2016) analyzed related items on home-based parental involvement included in PIRLS 2011 in a cross-cultural context. To do so, they validated the scale across different education systems, and found a negative relationship between home-based parental involvement and students' literacy achievement (Punter et al., 2016). South Africa was not among the educational systems included in the analyses. On that note, this study aims at analyzing whether the association of home-based parental involvement activities as administered in PIRLS with student outcomes remains negative for the South African context too.

Utilizing the applicability of PIRLS scales on (early) parental involvement activities, Caro (2018) and Feld (2018) analyzed thereof in association with student outcomes in learning across different educational systems. Caro (2018) looked at parental involvement in early literacy and numeracy activities in 28 education systems, and analyzed thereof with regard to students' subject specific interest (Caro, 2018). The latter is represented by two scales, namely whether students like learning mathematics and reading, and on their intrinsic motivation to learn. Furthermore, socioeconomic status is considered as a mediator for the association, and analyses are based on the TIMSS-PIRLS 2011 data set. Caro (2018) found that the mediated effect of early literacy activities performed by parents is negligible, thus the direct effect on students' interest in mathematics and reading is more relevant (Caro, 2018). Based thereon, Caro (2018) summarized that parental early learning activities serve as a potential "gateway for influencing student motivation across educational systems" (Caro, 2018, p. 16). Furthermore, comparing the Canadian and the German education systems, Feld (2018) found that parental support in children's schoolwork, that is helping to practice reading and mathematics, and parents' self-perceived invitations for school-based involvement, are associated with student achievement in reading in both educational systems.[31]

[31] On a side note, comparisons between educational systems are based on ensuring the invariance of the latent measures utilized in the study (Feld, 2018).

4.2.3. Parental Involvement in Association with Student Outcomes

As touched briefly throughout the state of research, some parental involvement activities are considered more influential than others. However, the effect may differ with regard to different forms of student outcomes in learning (W. Fan & Williams, 2010).
Besides being predictive of subject-specific achievement such as literacy performance, parental involvement is also known to be positively associated with students' personal motivational beliefs such as students' intrinsic motivation and self-efficacy (Eccles et al., 1998; W. Fan & Williams, 2010; Walker et al., 2005). Along the lines of the aforementioned study by Caro (2018), Fan & Williams (2010) assessed parental involvement activities with regard to tenth-grade students, and found that parental advising (e.g., selecting course programs at school) is positively associated with students' academic self-efficacy in English, and with the intrinsic motivation to learn (W. Fan & Williams, 2010).

Recent findings obtained from structural equation models based on PIRLS 2011 focusing the effect of students' self-efficacy and students' intrinsic motivation on overall reading achievement in Hong Kong have been presented by Cheung et al. (2017). Fourth-grade students' self-perceived motivation to read and reading self-efficacy predicted by parental early literacy activities showed a direct association with overall reading achievement. Though, the effect on self-efficacy appears to be stronger (.38) than the effect on reading motivation (.09). In total, the model explained 34% of the variance in reading achievement (W. M. Cheung et al., 2017). With their study, the authors addressed a research gap that investigated the effects of students' personal motivational beliefs on literacy proficiency in a non-western context and with regard to a non-alphabetical language such as Chinese (W. M. Cheung et al., 2017).
Utilizing PIRLS 2011 data for fourth-grade students in Abu Dhabi, Yang et al. (2018) used regression models to analyze the association between multiple home-background factors, namely parental involvement and students' personal motivational beliefs (i.e., motivation to read and reading self-efficacy) with students' reading achievement. In alignment with other studies, the authors reinforced that students' self-efficacy functions as the strongest

predictor for student reading achievement (Yang et al., 2018). Interestingly, Yang et al. (2018) also found that parental involvement activities have a significantly negative effect on reading achievement. Furthermore, the motivation to read is not significantly associated with reading achievement (Yang et al., 2018). This is despite the fact that, on an international average, analyses based on PIRLS 2011 data showed a positive association between students' reading motivation and reading achievement (I. V. S. Mullis, Martin, Foy, & Drucker, 2012). Hence, Yang et al. (2018) argue that the insignificant effect of motivation to read on students' reading achievement in Abu Dhabi might be due to the fact that students of younger ages "tend to overrate their reading beliefs and competences" (Yang et al., 2018, pp. 12–13).

Lastly, a recent publication by Chen et al. (2021) made use of PIRLS 2016 data for fourth-grade students in Canada, and investigated the association between student, home, and school factors with student performance in reading. In this study, the authors drew on self-determination theory, and implemented self-efficacy as a mediator for the association of aforementioned context factors and reading achievement. Results of the multilevel structural equation model showed a predominantly positive effect of predictor variables such as early literacy involvement of parents on students' self-efficacy in reading; and student self-efficacy on students' overall performance in reading (F. Chen et al., 2021).

4.2.4. Parental Involvement in the light of Socioeconomic Status

Taking into account the considerations of EV-Theory as well as cost-benefit calculations (sec. 3.2.2), educational investments such as parental involvement and the motivation for it have been found to be strongly tied to socioeconomic status. For example, according to the revised Model of Parental Involvement Processes, motives for parental involvement are parents' perceived life context, including time and energy as well as skills and knowledge of parents (Walker et al., 2005). Assuming that parents of lower socioeconomic status are more likely to lack those, the extent of their ability to offer involvement activities is often diminished (e.g., Bempechat & Shernoff, 2012; Wigfield et al., 2015). Thus, as shown empirically and conceptualized theoretically, socioeconomic status is con-

sidered an indispensable factor for the association of parental involvement with students' educational attainment. Therefore, it is unsurprising that most studies account for socioeconomic status in one way or another when analyzing the association between the two.

According to the perspective of socioeconomic status being associated with educational choices and educational attainment (Boudon, 1974; Esser, 1999), some studies showed that parental involvement is a leading mediator for this association (Grolnick & Slowiaczek, 1994). For example, studying cross-sectional data from the U.S., Stevenson and Baker (1987) found that the association between parental educational status, functioning as a proxy of socioeconomic status, and student achievement is strongly mediated by parental involvement in school activities (Stevenson & Baker, 1987). Furthermore, Epstein (1987) summarized that parental involvement forms, both at home and at school, positively "affect children's achievement, attitudes and aspirations, even after student ability and family socioeconomic status are taken into account" (Epstein, 1987, p. 120). Similar effects of socioeconomic status were confirmed by Lavenda (2011) and Caro (2018). For example, Caro (2018) found that the effect of early childhood activities performed by the parents showed a strong association with students' personal motivational beliefs to learn, yet, the mediating role of socioeconomic status is negligible (Caro, 2018).

Important work on the interplay of socioeconomic status and parental involvement was presented by Lareau (1987, 2000, 2002, 2003). Exploring the effect of social status on parental involvement activities of European American parents in the U.S. from a qualitative perspective, Lareau (1987) showed that respective activities of parental involvement are determined by socioeconomic status. For example, parents of different socioeconomic status, here working-class versus middle-class parents, act differently on invitations for involvement from the school, although they all attempt for their children to be educationally successful. Lareau (1987) assumes that this is due to the fact that working-class parents perceive themselves as less able to support their children in schoolwork, and therefore rather refrain from getting involved (Lareau, 1987). Extending her research, Lareau (2002) showed furthermore that socioeconomic status also outweighs ethnicity (Lareau,

2002). Mirroring the premises of EV-Theory, respective educational non-investments of working-class parents are likely to result in their children remaining in lower social classes, and incidentally further support that "the apple doesn't fall far from the tree" (Esser, 1999, p. 265).

Taken together, results on the interplay of socioeconomic status and parental involvement, student outcomes in learning, and the association thereof, are mixed. Whereas some of the aforementioned studies ascribe a positive effect of socioeconomic status (e.g., Lareau, 1987, 2002; Stevenson & Baker, 1987), others point towards a negligible effect thereof (e.g., Caro, 2018), or the existence of an association between parental involvement and student outcomes in learning despite controlling for socioeconomic status (e.g., Epstein, 1987; Lavenda, 2011). Not least due to the indecisiveness of its effect, this study accounts for socioeconomic status by considering two socioeconomically different groups for the analysis of parental educational investment behavior and parental involvement.

4.2.5. Parental Involvement in Cultural Contexts

The subject of parental involvement has been highlighted along the lines of motivational and volitional activities, as well as their association with student outcomes in learning. However, along the lines of Walker et al. (2011), who pointed towards a gap in understanding context-specific patterns of motivation for parental involvement activities (Walker et al., 2011), researchers acknowledged differences in the association of parental involvement activities and educational outcomes peculiar to cultural contexts (e.g., Chiu, 2010). Hence, the following section focuses parental involvement activities in different cultural contexts, emphasizing on South Africa.

International Findings

Evidently, studies predominantly looked at parental involvement in Europe, Asia, the U.S., and Canada (Punter et al., 2016). Thus, respective studies consider different ethnicities such as Latino, African American, or Black American (e.g., Jeynes, 2003, 2007;

Lareau, 1987, 2002; Lavenda, 2011; Peng & Wright, 1994; Seyfried & Chung, 2002; Walker et al., 2011), as well as different language groups (e.g., Yang et al., 2018).

For instance, Peng & Wright (1994) focused parents of different ethnic groups in the U.S. with regard to their involvement activities and its effect on student outcomes in learning. They found that the home environment as well as home-based educational activities of students and parents are associated with differences in student achievement between ethnic groups. Hence, they conclude that this "finding is valuable because it provides a basis for developing strategies for narrowing the differences in achievement among racial-ethnic groups" (Peng & Wright, 1994, p. 351). However, Lavenda (2011) stated, that the Model of Parental Involvement Process after Hoover-Dempsey & Sandler (1995, 1997) has not been tested in non-U.S. contexts, and thus, utilized it to test school-based parental involvement for Jewish and Arab parents in Israel (Lavenda, 2011). With regard to cultural differences, Lavenda showed that a rather small share of variance in parental involvement was explained by ethnicity (Lavenda, 2011).

Also utilizing data from the U.S., Seyfried & Chung (2002) presented a study that distinguished parental involvement activities of African American and European American eighth-grade students. With regard to group differences in parental involvement, they found that although African American families have similarly high educational aspirations for their children, respective attitudes do not translate into actions that foster significant changes of student outcomes in learning. The opposite is true for the comparison group, which are European Americans (Seyfried & Chung, 2002). Respective findings are in line with those presented by Lareau (1987, 2002), claiming that educational aspirations do not differ between ethnicities, yet quality of parental involvement does (Lareau, 1987, 2002).

Findings on the South African Context

Taken together, there are still very little results on the subject of parental involvement for the South African context. In fact, only few and mainly qualitative studies, were identified.

For example, early pioneering work has been conducted by van Wyk (1996), who presented a qualitative study investigating parental involvement activities among Black-African parents in township schools. The author found that Black-African parents faced several issues that hinder involvement activities. Among those were illiteracy, lack of proficiency in language of instruction, and poverty. Furthermore, van Wyk (1996) showed that students faced behavioral problems such as lack of discipline and lack of intrinsic motivation to learn (van Wyk, 1996).

Inspired by characteristics of parents allegedly impeding involvement activities, Bojuwoye & Narain (2008) focused on parental educational aspirations as well as home- and school-based involvement activities in association with student academic achievement in South African high-schools. Volitional parental involvement comprised of typical activities such as involvement in homework, and support in subject-related practices such as reading. It was found that students of parents who claimed to be highly involved, performed better than their counterparts (Bojuwoye & Narain, 2008). However, Bojuwoye & Narain (2008) based their findings on descriptive statistics such as correlation analysis and did not account for differences between language groups and socioeconomic status.

Lastly, a qualitative study by Singh et al. (2004) focused the involvement activities of Black-African parents in South Africa, concluding that the effect of parental involvement, especially in terms of homework support, is decisive for students' educational performance (P. Singh et al., 2004). However, the authors also emphasized the importance of socioeconomic status. In line with findings by Lareau (1987, 2002), Singh et al. (2004) confirmed that parents of lower socioeconomic status perceive themselves as incapable to support children's academic achievement, and hence, relied on the teacher to do so (P. Singh et al., 2004). Along those lines, Felix et al. (2008) presented a qualitative intake on the subject, and went as far as stating that "parents from poorer socioeconomic groups [in South Africa] were positioned as disinterested and unable to assist their children" (Felix et al., 2008, p. 99).

In addition to the aforementioned results based on qualitative results, van Staden & Howie

(2012) utilized data from PIRLS 2006, and showed that opportunities to read created by South African parents had a positive effect on reading achievement of fifth-grade students. However, parents spending time to read with their children had an insignificant effect. Furthermore, students' socioeconomic background showed to be predictive of reading performance as well (van Staden & Howie, 2012). In a more recent approach, Combrinck et al. (2014) utilized data from prePIRLS 2011 and investigated the effect of parental involvement activities in early childhood activities with regard to fourth-grade students' performance in reading. Results pointed towards the importance of early learning activities for academic achievement in reading. With few notable exceptions, the authors found that regular parental involvement in early childhood education, comprising activities such as frequently telling stories, was highly beneficial to the overall reading performance of fourth-grade students in South Africa (Combrinck et al., 2014). The subsequent question is whether the effect remains valid when socioeconomic status and language groups are accounted for.

On that note, van Staden & Howie (2014) analyzed a set of factors predictive to educational performance in reading of fifth-grade students in South Africa. Among others, the authors accounted for different language groups as well as students' socioeconomic background. They obtained notable effects of individual student characteristics, including the frequency of parental involvement activities at home. For instance, activities fostering reading literacy showed positive associations with learning outcomes in the language group *Afrikaans*, and negative associations in the language groups *English* and *Sotho* (van Staden & Howie, 2014).

4.3. Summary and Identification of Research Gaps

The goal of this section was to provide a general overview on the state of research regarding educational investments as well as parental involvement as a form thereof, and to identify relevant research gaps. For this purpose, results on the subject from both national and international perspectives were summarized with regard to versions of student outcomes in learning and socioeconomic status under consideration of different ethnic backgrounds.

Summary

The first step was to take a look at educational investments as a potential driver for educational inequalities. Predominantly, studies that put EV-Theory to the test were summarized. Given that the frameworks were primarily developed based on educational systems of the U.S. and Europe, the body of research comprises of many studies focusing on respective countries. However, Ardington et al. (2011) and Lam et al. (2011) contributed important work for the South African context and in regard to the effect of educational investments on differences in educational attainment. Doing so, they acknowledged different societal groups, addressing the effect of socioeconomic status and different ethnicities, for instance Black-Africans and Whites. Taken together, educational investments may well be seen as a driver of educational inequality, given that empirical research was able to confirm that the ability to take educational investments depends on individual resources such as time and monetary means, and socioeconomic status.

The second step was to explore parental involvement as a form of educational investment and highlight its effect on student outcomes in learning under consideration of socioeconomic status and with regard to different cultural contexts. Along the lines of the theoretical framework, motivational and volitional aspects of parental involvement activities were considered. Given that this study focuses home-based parental involvement, the topic was emphasized in this section.
The body of research showed indecisive results on the question whether parental involvement activities are beneficial to student outcomes in learning and in regard to identifying

activities that are most effective. However, invitations for involvement from others as perceived by the parents, and parental aspirations for their children's education were generally found to have a positive effect on parental involvement activities. Compared to single parental involvement activities, a global indicator of parental involvement was found to be more predictive of student educational attainment. With regard to the latter, research showed that parental involvement is generally positively associated with students' personal motivational beliefs, such as intrinsic motivation to learn and self-efficacy.

Lastly, this section addressed the effect of socioeconomic status on parental involvement, and looked at involvement activities in different cultural contexts. The effect of socioeconomic status has been assessed differently across various studies, resulting in mixed conclusions regarding its relevance to the association of parental involvement and student outcomes in learning. However, results for the South African context showed that socioeconomic status may well be considered an influential predictor in the equation of parental educational investment behavior, that is parental involvement.

Identification of Research Gaps

Based on the current state of research summarized and reflected in this section, the following research gaps are identified for this study.

First, what is missing in the discourse of educational inequality and its persistence in post-Apartheid South Africa is the application of theoretical concepts to systematically explore and describe mechanisms that feed dysfunctional societal patterns. This is not least due to a shortcoming of theoretical concepts that are tailored to explaining respective mechanisms in countries such as South Africa (van Staden & Howie, 2014, p. 172). Although Ardington et al. (2011), and Lam et al. (2011) utilized Human Capital Theory in order to explain educational decision making in South Africa, they argued from an economical perspective. With an exception by van Staden & Howie (2014), who provided empirical work based on Creemers' Comprehensive Model of Educational Effectiveness, sociological

concepts have not yet been vastly applied to the context of South Africa's educational system. By utilizing a framework that is built on different elements of psychological and sociological Rational Choice Theory, this study aims at counteracting respective shortcomings, and thus, support research on the applicability of theoretical frameworks for the sufficient investigation of educational investment behavior in South Africa.

Second, with regard to parental involvement as a form of educational investment behavior, this study aims at utilizing the revised Model of Parental Involvement Process by Walker et al. (2005), and address the lack of investigating its applicability in different cultural contexts as claimed by Fan & Chen (2001). That is, apart from work shown by Howie (2003), van Staden & Howie (2012, 2014), and Combrinck et al. (2014), most of the results on parental involvement in South Africa are based on small-scale and mostly qualitative studies. The data set of prePIRLS offers an extensive set of assessment and background data, allowing for representative assumptions on the association between parental motivation and volition with student outcomes in learning. Furthermore, aimed at accounting for racial segregation in South Africa, respective mechanisms are investigated for groups of the South African population. This is in line with work by van Staden & Howie (2014), who controlled for language groups, as well as studies that based the differentiation of groups on ethnicity and socioeconomic status (e.g., Felix et al., 2008; P. Singh et al., 2004; van Wyk, 1996).

Third, with regard to parental involvement as a form of educational investment behavior, a review of relevant meta-studies and individual research projects showed that motivational and volitional dimensions of parental involvement, as well as their association with student educational attainment, have scarcely been analyzed empirically (e.g., X. Fan & Chen, 2001). That is, studies primarily operationalized the revised Model of Parental Involvement Process by Walker et al. (2005), and focused on why and how parents get involved (e.g., Green et al., 2007; Reiniger & López, 2017; Walker et al., 2011). A particular case in that regard is the aspect of parental educational aspirations. Defined as a form of parental involvement, researchers usually claim parental aspirations to be an active behavior that is most often positively associated with student attainment (e.g.,

X. Fan & Chen, 2001; Peng & Wright, 1994; Seyfried & Chung, 2002; Xu et al., 2010). However, in the light of models considering linearity of human actional behavior (e.g., the Rubicon Model of Action Phases after Heckhausen & Gollwitzer (1987), and the revised Model of Parental Involvement Process after Walker et al. (2005)) as a blueprint for mapping actional behavior, the understanding of parental educational aspirations as a dimension of volitional behavior is inaccurate. More precisely, it must be considered an attitude with the potential to motivate certain activities of parental involvement (e.g., K. Singh et al., 1995). To that end, educational aspirations are understood as a motive for involvement activities, rather than a volitional activity itself. Accordingly, the association between parental educational aspirations and student outcomes in learning is not modeled directly, but through parental involvement. This approach is in line with early work by Singh et al. (1995).

To conclude, implications for the study drawn from the current state of research include the need for the investigation of educational investment behavior utilizing theories of rational decision making from a non-economical perspective, and a transparent operationalization of parental involvement as well as the interpretation of its effects on student outcomes in learning for the South African context. This is not least because multiple definitions and conceptualizations of parental involvement led to inconsistencies of effects. The respective particularity is considered when interpreting the results of this study and positioning them within the state of research.

5. Research Questions & Hypotheses

The following section phrases research questions as well as corresponding hypotheses, based on the theoretical framework and state of research presented in the preceding sections 3 and 4 of this study.

5.1. Research Questions

Aiming to add to the understanding of mechanisms underlying persisting (educational) inequalities in South Africa, the main focus of this study is on the educational investment behavior of South Africans, which are based on differences in socioeconomic status, and partly account for the parameters identified in section 2. To this study, socioeconomic low-performers are of particular interest. They form the focus group, and thus, the unit of reference for the following research questions and hypotheses.

Throughout the study, the reference group is explored from two angles and along the lines of the theoretical framework (sec. 3). To approach the analyses of educational investment behavior and parental involvement, an adapted model of human action behavior is utilized (fig. 6). A special characteristic of the model is the causal chain of elements, that is, motivation is followed by volition, and volition is followed by the evaluation thereof. Thus, the former naturally conditions the latter. This study picks up on that as well, and attempts to mirror the associations between phases of motivation, volition and evaluation, respectively.

Along those lines, the following research questions are applied:

I. Are socioeconomic low-performers in South Africa motivated to invest in education?

II. Are socioeconomic low-performers in South Africa transferring motivation into (parental) educational investments?

III. Are parental educational investments of socioeconomic low-performers in South Africa beneficial to student outcomes in learning?

The qualitative perspective of this study explores whether socioeconomic low-performing street vendors in South Africa are motivated to invest in education, and whether they transfer respective motivation into actions. This part of the study serves as a door opener, highlighting general attitudes towards educational investments and elaborating on different types thereof. Thus, the qualitative perspective aims at answering research question I and II.
The quantitative perspective is conducted sequentially to the qualitative perspective. It is focused on parental involvement as a specific form of educational investments. In this part, motivation for and actions of parental involvement are analyzed with regard to student outcomes in learning. Thus, the quantitative perspective aims at answering research question I, II and III.

Finally, this study aims at contributing to the understanding of educational investment behavior, and at identifying crossing points, where lack of socioeconomic resources might limit the pursuit of parental educational investments. Based thereon, this study endeavors to provide empirically supported suggestions for policies that potentially counteract the persistence of educational inequalities in post-Apartheid South Africa.

5.2. Hypotheses

Hypotheses of this study are presented next. If applicable, hypotheses are accompanied by the complementary null hypothesis (e.g., $H_{1.0}$). Given that this study is primarily interested in (parental) educational investment behavior of South African socioeconomic low-performers, this group is identified as the focus group (i.e., the reference group). Thus, the following hypotheses evolve around it. Throughout the quantitative study, the reference group is compared to counteracting groups, aiming at putting potential patterns of educational investment into perspective. This is reflected in hypotheses 3 to

9. Hypotheses 1 and 2 apply to both, the quantitative and the qualitative perspective.[32]

5.2.1. Motivation and Actions towards Educational Investments

Building on the perspective of rational decision making represented by EV-Theory, and associated principles of cost-benefit calculations, it is expected that socioeconomic low-performers are neither motivated to invest in education nor to take educational investments. That is primarily because of the expectations, namely the costs required for educational investments, and the uncertainty of educational outcomes to be successful. Aside from the direct costs, time is identified to be a relevant factor as well, given that the attainment of a higher educational degree requires more time, and thus, accumulated opportunity costs. EV-Theory assumes that costs and associated risks are likely to be taken when the motivation for an investment exceeds them. This is often observed among socioeconomic high-performers when status preservation is at stake (Esser, 1999). Assuming a universal validity of the association between costs and benefits, respective behavioral patterns are expected to be found across different cultural contexts. However, recent empirical findings on the South African population indicate that parents of all socioeconomic backgrounds in South Africa have high educational aspirations for their children. Furthermore, studies show high parental involvement, for example, in early learning activities as well as daily activities of involvement in their children's schooling (e.g., Howie et al., 2012, 2018).

Taking into account respective principles of rational decision making as well as empirical evidence summarized from the state of research, the general contradictory hypotheses, applicable to both qualitative and quantitative perspectives of this study, are the following.

[32] In the case of the qualitative study, hypotheses only serve as a reference point for analysis and interpretation of interview materials. Thus, the qualitative perspective does not imply hypotheses testing in a statistical sense.

$H_{1.0}$: Socioeconomic low-performers in South Africa are not motivated to invest in education.

$H_{1.1}$: Socioeconomic low-performers in South Africa are motivated to invest in education.

$H_{2.0}$: Socioeconomic low-performers in South Africa do not take investments in education.

$H_{2.1}$: Socioeconomic low-performers in South Africa take investments in education.

5.2.2. Motivation and Actions towards Parental Involvement

With regard to the motivational dimension and its association with parental involvement activities, effects of invitations for involvement from others, and parental educational aspirations are well researched with regard to parental involvement activities. The following has been found.

Invitations for involvement from others is one of the dimensions motivating parental involvement activities according to the revised Model of Parental Involvement Process after Walker et al. (2005). Overall, research confirmed the relevance of the dimension on parental involvement (e.g., Desland & Betrand, 2005). However, some studies came to a different conclusion, showing no significant effect of invitations for involvement on parental involvement activities (e.g., Green et al., 2007; Reiniger & López, 2017; Walker et al., 2011).

With regard to the effect of parental educational aspirations, results are inconclusive as well. That is despite the fact that parental educational aspirations are generally assumed to be of significant importance for motivating educational investments and parental involvement activities (e.g., X. Fan & Chen, 2001). Whereas various studies found a significant effect of parental educational aspirations on either parental involvement or student outcomes in learning (e.g., Peng & Wright, 1994; Seyfried & Chung, 2002; K. Singh et al., 1995; Xu et al., 2010), some studies found no noteworthy associations in that regard (e.g., Jodl et al., 2001).

Considering the effect of socioeconomic status on parental motivation for educational investments, research showed that parental involvement manifests differently between socioeconomic groups, even though parents of different groups have similar perceptions of invitations for involvement from the school and educational aspirations. Potential explanations are that parents of lower socioeconomic status perceive themselves as incapable of supporting their children in school (e.g., Lareau, 1987; Seyfried & Chung, 2002).
On that note, hypotheses addressing the effect of factors motivating parental educational involvement while considering socioeconomic status are the following.

$H_{3.0}$: Compared to socioeconomic high-performers, invitations for involvement from others are insignificantly associated with parental involvement among socioeconomic low-performers in South Africa.
$H_{3.1}$: Compared to socioeconomic high-performers, invitations for involvement from others are significantly associated with parental involvement among socioeconomic low-performers in South Africa.

$H_{4.0}$: Compared to socioeconomic high-performers, educational aspirations are insignificantly associated with parental involvement among socioeconomic low-performers in South Africa.
$H_{4.1}$: Compared to socioeconomic high-performers, educational aspirations are significantly associated with parental involvement among socioeconomic low-performers in South Africa.

With regard to the volitional dimension of home-based parental involvement, and its association with student outcomes in learning, the current state of research showed inconclusive results. Given that parental involvement is often inconsistently defined, effects on student outcomes in learning differ as well. Along those lines, positive effects of global parental involvement activities have been confirmed regarding student outcomes in learning (e.g., X. Fan & Chen, 2001; Jeynes, 2005, 2007), as well as students' personal

motivational beliefs (e.g., F. Chen et al., 2021; Eccles et al., 1998; W. Fan & Williams, 2010; Grolnick & Slowiaczek, 1994). However, concerning the latter, single studies also identified a negative effect of parental involvement activities (e.g., Yang et al., 2018).
In line with the assumptions of EV-Theory, research furthermore identified socioeconomic status to be predictive of parental involvement activities. Assuming that parents of lower socioeconomic status are more likely to lack resources required for involvement activities, for example time and energy as well as skills and knowledge (e.g., Lareau, 1987; Seyfried & Chung, 2002; Walker et al., 2005), the extent of their involvement activities is often diminished (e.g., Bempechat & Shernoff, 2012; Wigfield et al., 2015). Furthermore, the motive of status preservation as a driver for educational investments is less likely to function for socioeconomic low-performers (Esser, 1999). Though, some studies showed the negligible effect of socioeconomic status on parental involvement activities, and its association to student outcomes in learning altogether (e.g., Caro, 2018; Lavenda, 2011).
Hence, hypotheses addressing the effect of parental educational involvement activities on student outcomes in learning are the following.

$H_{5.0}$: Home-based parental involvement is negatively associated with student outcomes in learning.
$H_{5.1}$: Home-based parental involvement is positively associated with student outcomes in learning.

$H_{6.0}$: Compared to socioeconomic high-performers, home-based parental involvement is insignificantly associated with student reading achievement among socioeconomic low-performers in South Africa.
$H_{6.1}$: Compared to socioeconomic high-performers, home-based parental involvement is significantly associated with student reading achievement among socioeconomic low-performers in South Africa.

$H_{7.0}$: Compared to socioeconomic high-performers, home-based parental involvement is insignificantly associated with students' personal motivational beliefs among socioeconomic low-performers in South Africa.

$H_{7.1}$: Compared to socioeconomic high-performers, home-based parental involvement is significantly associated with students' personal motivational beliefs among socioeconomic low-performers in South Africa.

5.2.3. Associations of Additional Concepts

Lastly, hypotheses capturing the associations between dimensions of parental involvement, and between students' personal motivational beliefs are stated.

In general, parental involvement is considered multidimensional (e.g., X. Fan & Chen, 2001; Walker et al., 2005), yet some argue for the need to consider related activities separately rather than applying an index of the construct (X. Fan & Chen, 2001). However, this study opts for a multidimensional operationalization of home-based parental involvement, yielding in parental activities of *Supervision* and *Practice* (sec. 8.3). By doing so, this study aims at depicting more precise and contextually cohesive measures of parental involvement activities. However, assuming that they somewhat overlap as well, a correlation of the two is assumed.

Furthermore, outcome variables, namely student reading achievement and students' personal motivational beliefs (i.e., intrinsic motivation and self-efficacy in reading), are considered to be associated (e.g., Bong, 2008; W. Fan & Williams, 2010). On that note, it is assumed that students' self-efficacy is more influential to students' overall performance in reading than students' intrinsic motivation to learn (e.g., W. M. Cheung et al., 2017).

Finally, hypotheses addressing the association between dimensions of home-based parental involvement, and associations between student outcomes in learning are the following.

$H_{8.0}$: Dimensions of home-based parental involvement are negatively associated with each other.
$H_{8.1}$: Dimensions of home-based parental involvement are positively associated with each other.

$H_{9.0}$: Students' reading achievement and students' personal motivational beliefs are negatively associated with each other.
$H_{9.1}$: Students' reading achievement and students' personal motivational beliefs are positively associated with each other.

6. Study Design

In order to explore educational investment behavior among socioeconomic low-performers in South Africa, this study comprises of a quantitative study, initiated by the results of a prefixed qualitative study.

6.1. Combination of Qualitative and Quantitative Studies

The qualitative study (sec. 7) was conducted in 2013 and looked at South African street vendors' investment behavior towards their own educational career. To that end, 12 guided interviews with street vendors selling the *The Big Issue* South Africa (TBI) in Cape Town/South Africa were carried out. The data was analyzed along the lines of empirically grounded construction of types and typologies (Kelle & Kluge, 1999, 2010). Subsequent to the qualitative study is the quantitative study (sec. 8). Utilizing South African data from prePIRLS 2011, the quantitative study focuses the analysis of parental involvement as a form of educational investment behavior, and its association with student outcomes in learning. For that purpose, multi-group structural equation models are analyzed. Figure 7 shows the theoretical framework of this study, and appoints the scope of the qualitative and the quantitative study.

The qualitative study focuses the pre-actional and volitional phase of human action behavior, emphasizing general educational investment behavior of socioeconomic low-performs towards their own educational career. This is indicated by the dashed circle. The quantitative study focuses the association between motivation, volition, and outcomes of parental educational investment behavior with regard to their children's educational attainment. This is indicated by the solid circle.

6.2. Methodological Considerations on Mixed-Methods Designs

Although shedding light onto the educational investment behavior of socioeconomic low-performers from different perspectives, both studies are evaluated with regard to the same theoretical framework, linking them contextually rather than empirically. On that note,

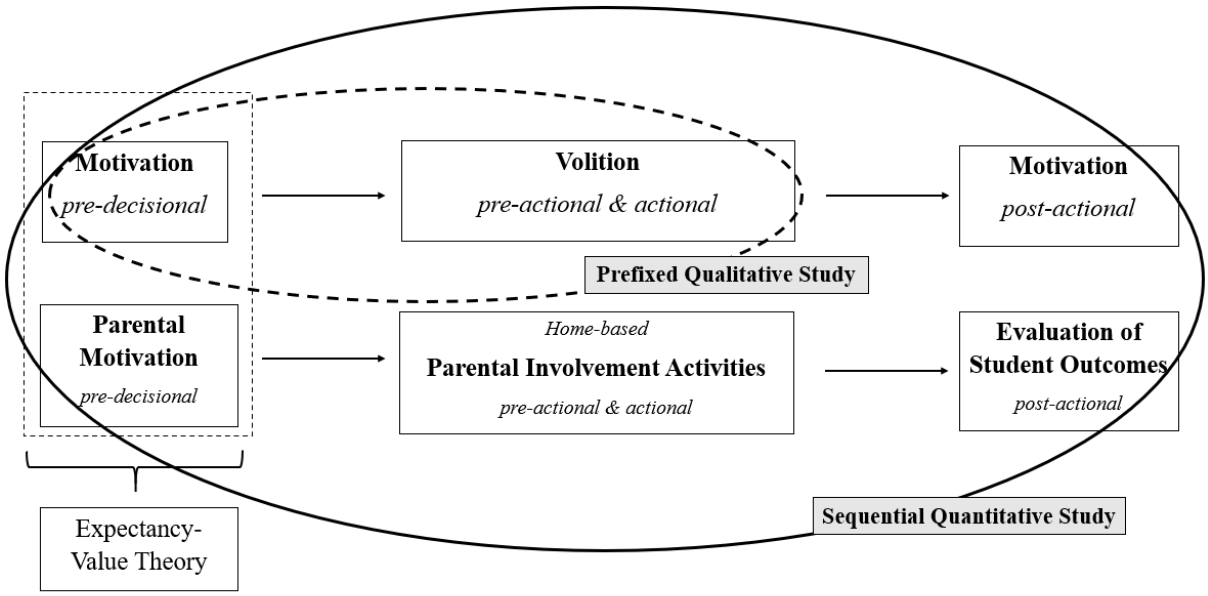

Figure 7: The methodological study design, based on the adapted Rubicon Model of Action Phases (J. Heckhausen & Heckhausen, 2008) and the (revised) Model of Parental Involvement Process (Walker et al., 2005), substantiated with theories of rational decision making. Own illustration, appointing the scope of the qualitative and quantitative perspective.

this study is based on a two-phase sequence of a qualitative and a quantitative study, with the qualitative study indicating the need of exploring the subject of interest further. However, although it seems supposedly appropriate at first, the layout of this study does not meet the premises of multiple- or mixed-methods design. That is due to the following reasons.

Up to this point, the topic of mixed-methods research has been broadly discussed, especially with regard to the question of what characterizes it. Earliest attempts on the "third methodological movement" (Johnson et al., 2007, p. 118) are based on work by Campbell and Fiske (1959), and ever since, multifarious names and definitions of mixed-methods research are found throughout the literature. However, the most common terminology used today is multiple-methods research, and mixed-methods research as a subdomain thereof, specifically addressing the combination of qualitative and quantitative research in one larger setting of a study (Fetters & Molina-Azorin, 2017).

A prominent definition of mixed-methods design was given by Johnson & Onwuegbuzie (2004, 2007), who referred to it as a methodological approach where qualitative and quantitative "research techniques, methods, approaches, concepts or language" are mixed into a single study (Johnson & Onwuegbuzie, 2004, p. 17). Thus, mixed-methods designs "legitimate the use of multiple approaches in answering a research question" (Johnson & Onwuegbuzie, 2004, p. 17). In terms of a more comprehensive definition of mixed-methods, Creswell & Plano Clark (2007) stated the following:

> "As a method, it focuses collecting, analyzing, and mixing both quantitative and qualitative data in a single study or series of studies. Its central premise is that the use of quantitative and qualitative approaches in combination provides a better understanding of research problems than either approach alone." (Creswell & Plano Clark, 2007, p. 5)

Furthermore, researchers explicitly point to the flexibility of mixed-methods research, and thus, enable practical design decisions (e.g., Creswell & Plano Clark, 2007; Schreier & Odağ, 2010). However, the decisive factor for the identification of mixed-methods research is always whether data is sufficiently *mixed*. Mixing can be achieved by merging,

embedding or connecting qualitative and quantitative data (Creswell & Plano Clark, 2007). On that note, “it is not enough to simply collect and analyze qualitative and quantitative data; they need to be "mixed" in some way so that together they form a more complete picture of the problem than they do when standing alone” (Creswell & Plano Clark, 2007, p. 7).

Although this study certainly uses a mix of data and methods to approach the research question of interest from different perspectives, it is not considered a mixed-methods design by definition. That is, in the sense of mixing both perspectives in a way that they are joined in the attempt of answering the research question of interest. Hence, the qualitative and the quantitative study are presented and interpreted separately. A careful consolidation of results with regard to educational investment behavior of socioeconomic low-performers in South Africa is carried out during the discussion (sec. 9).

7. Qualitative Study: General Educational Investments

Data, methods, and results of the qualitative study are presented next to lay the grounds for the sequentially conducted quantitative study.[33] Since there is no official documentation available for the qualitative data set and in order to ensure sufficient transparency of research practice, special emphasis is put on the description of the sample, methodology of data collection, as well as conceptual limitations of the study first. Next, methods for data analyses are elaborated, and results of the study are presented.

As appointed throughout this section, the qualitative data set was originally collected to elaborate educational needs of street vendors in Cape Town/South Africa were at the time of the study. However, data of the original study beared the potential for qualitative secondary analysis (e.g., Heaton, 2008; Medjedovic, 2008; Witzel et al., 2008), and hence, has been reused in order to explore educational investment behavior of street vendors. Given that the secondary qualitative study has been built on self-collected data (Heaton, 2008), full control over the data collection and processing has been ensured.

7.1. Data

The basis for the qualitative analyses is a study from 2013, which was developed and carried out during a three-month research internship in Cape Town/South Africa. Conducting guided interviews, the study aimed at exploring the educational needs of street vendors in Cape Town/South Africa.

The research internship was supported by a scholarly travel grant provided by Engagement Global, and hosted by the *DVV International Project Office South Africa*,[34] and

[33] Data and results of this study as presented in this section have been part of my unpublished master thesis (Twele (formerly Radermacher), 2015) to obtain a graduate degree in Sociology and Social Research at the University of Bremen in 2015.

[34] DVV International is the *Institute for International Cooperation* of the *German Adult Education Association*. DVV International is actively involved in promoting and supporting youth and adult education around the world, aiming at ensuring life-long learning opportunities for all, and thus, works along the lines of the United Nations' *Sustainable Development Goals*. DVV International has been represented in South Africa for 20 years. Further information are available at: www.dvv-international.de

The Big Issue South Africa (TBI).[35]

7.1.1. Sampling Procedure

The sample of the qualitative study consists of South African street vendors selling TBI as informally working entrepreneurs in Cape Town/South Africa and its immediate vicinity. As presented in section 2.2.3, the informal sector is particularly important for South Africa. Hence, in 2018, it made about 30% of the country's total work (Rogan, 2019).
Field access was facilitated by TBI. Employees of the magazine selected street vendors to be interviewed, and provided premises in the main office located in Cape Town/South Africa. Hence, the sample resulted from a purposeful sampling approach aimed at homogeneously sampling individuals based on their belonging to the group of street vendors for the purpose of investigating their specific educational needs (Creswell, 2016, p. 208). That being said, it was ensured that "the essential socio-structural context conditions relevant to the field investigated were taken into account when selecting research units" (Kelle & Kluge, 2010, p. 50). The sampling approach is considered a convenience sample. That is, because certain prerequisites needed to be met. For example, all interviewees needed to be able to master an interview in English, given that scarce resources did not allow for a translator to conduct and translate interview material in each interviewees' native language into English. Furthermore, personnel resources limited the sample size to n = 12.
As an incentive to take part in the interviews, interviewees were reimbursed for their

[35] Founded in 1997 and located in Cape Town/South Africa, *The Big Issue* South Africa (TBI) is an organization aiming to counteract local unemployment by pursuing a twofold approach. On the one hand, the potential of the informal sector (sec. 2.2.3) is used to provide income opportunities. On the other hand, informal learning opportunities are provided. The concept of TBI is straightforward. The organization publishes a monthly magazine, and offers it at a low price to the vendors, who are mostly socioeconomic low-performers such as marginalized and homeless people. The vendors, for their part, become independent entrepreneurs by selling the magazine for about 50% more than their initial purchase. Thus, the difference is the street vendors' profit. Thereby, TBI enables a direct source of income, as well as the possibility to learn independent and efficient business management at a small scale. In addition to providing the magazine, the organization offers a continuing educational program with workshops and training sessions on various topics. Further information are available at: www.bigissue.org.za

travel expenses and received a free copy of the magazine's current issue.

7.1.2. Description of the Sample

Demographic information gathered through a standardized questionnaire administered prior to the guided interview showed that interviewed street vendors (n = 12) were between 27 - 58 years old. Seven male and five female interviewees participated. 50% were single, but lived in households with up to five other residents.

More than half of the interviewees had no additional gainful employment other than selling TBI. In case of additional occupations, these were mostly in the domestic sector such as household services or gardening, as well as activities related to arts and crafts, and low-threshold employments as cashiers, security staff, or car washers. The reported average weekly income was ZAR390 ($23) considering all sources of income.[36] Thus, the daily income equaled ZAR55 ($3). Recall, the poverty line of upper middle income countries such as South Africa is estimated at $5,50 per day (sec. 2.2.3). Interviewed street vendors of this study earned below that, and are therefore considered socioeconomic low-performers.

Furthermore, interviewees indicated that although their working language is English, their primary language spoken at home is isiXhosa, which is the second most frequently spoken native African language in South Africa (16% in 2011), and the second most often used language of Black-Africans in South Africa (20% in 2011, 2.2.1).

Lastly, the average educational status of the interviewees was reported to be 10 years of schooling, or the completion of secondary education, respectively. One interviewee reported to have completed eight years of schooling. Two interviewees completed parts of *Further Education and Training*, that is post-secondary education in grades 10 - 12. Only one interviewee reported to have completed the *National Senior Certificate*, that is Matriculation, after 12 years of schooling. Hence, the majority of interviewees completed

[36] The average weekly income is based on the income information of 11 valid answers obtained from the interviewees. The mode of the income distribution is ZAR500 ($29). The minimum income was reported to be ZAR100 ($6) and the maximum income was reported to be ZAR500 ($29), resulting in a range of ZAR400 ($24).

the required level of compulsory education (sec. 2.2.4).

7.1.3. Methods of Data Collection

The Method of the Problem-Centered Interview

In order to gain a subjective insight to the educational needs of street vendors in Cape Town/South Africa, a qualitative approach was chosen (Flick, 2012). The method of the problem-centered interview after Witzel (1985, 2000) was then selected to ensure subject-specific information on the topic of interest. Based on the premises of Grounded Theory after Glaser & Strauss (1998), the method of problem-centered interviews operates inductively as well as deductively, and attempts to eliminate the contrast between strictly theory-guided attempts and completely open interview formats. The overall goal of the method is to objectively collect information on human behavior (Witzel, 2000).

The method of problem-centered interviews is characterized by three criteria, namely problem-centered orientation, process orientation, and object orientation (Witzel, 2000). First, problem-centered orientation describes the approach of orienting the research question towards a socially relevant problem. The interviewer makes use of the objective framework condition, or the context of action, which presuppose an individual action and which are not expected to change, aiming to understand the interviewee's explications. Furthermore, the interviewer is also encouraged to continue asking problem-oriented questions (Witzel, 2000). Second, process orientation refers to the principle of a sensitive and accepting communication process, aimed at building trust on the part of the interviewee throughout the research process. According to Witzel (1985, 2000), trust generates openness and enhances the ability to remember, as well as the motivation for self-reflection (Witzel, 2000). And third, object orientation accounts for the flexibility of the method towards different and individual requirements of the object that is being studied (Witzel, 2000).

The level of flexibility required for problem-centered interviews is achieved through a

combination of different tools, for example, group discussions, the biographical interview, and the qualitative interview (Flick, 2012; Witzel, 2000). According to Witzel (2000), the qualitative interview is the most important instrument. It consists of four elements, which are a short questionnaire, guidelines for the interview, audiotape recordings, and postscripts (Witzel, 2000).

First, a short questionnaire typically consists of a fixed question-answer scheme. It is applied prior to the interview and serves as a tool to collect contextual information, for instance on demographic information of the interviewees. Thus, the short questionnaire also enables the interviewer to center the guided interview around a-priori information relevant to the topic of interest (Witzel, 2000).

Second, guidelines for the interview have two major advantages. First, an applied interview guideline supports the interviewer in structuring the interview. Second, an interview conducted along the lines of a guideline enhances the comparability of information collected during individual interviews. Despite problem-oriented keywords or notes, an interview guideline typically consists of a preformulated introduction to the interview, general explorations of the problem, and ad-hoc questions (Witzel, 2000). The introduction to the interview as well as general explorations of the problem can serve to guide, deepen, and comprehend the content of the conversation during the interview (Flick, 2012; Witzel, 2000). Ad-hoc questions may contribute to ensuring the comparability of individual interviews. They are used, for instance, when a relevant aspect of a problem is excluded by the interviewee (Witzel, 2000)

Third, tape recordings are considered a standard tool in the field of qualitative interviews (Flick, 2012). They are used to facilitate the situation for the interviewer, and support the percipience of events occurring during the interview (Witzel, 2000).

Fourth, postscripts function as a written documentation of the conversation. Typically prepared immediately after the interview by the interviewer, they serve as a subjective, sketch-like documentation of verbalized contents, situational and non-verbal actions, as well as focal points addressed by the interviewee. Furthermore, postscripts aim at capturing conspicuous events noticed by the interviewer, such as body language or tone of

voice. Thus, postscripts function as a supplement to tape recordings (Witzel, 2000).

Implementation of the Problem-Centered Interview

The problem-centered interview was conducted as follows.

Due to scarce resources in terms of time, space, and availability of interviewees, the combination of different tools such as the qualitative interview, group discussions, and the biographical interview (Flick, 2012; Witzel, 2000) was discarded. Instead, only the qualitative interview was utilized for this study. Given that the qualitative interview is considered the most important instrument of the problem-centered interview, the correctness of the methodological approach is not diminished. However, the aforementioned elements of the qualitative interview, namely the short questionnaire, guidelines for the interview, tape recordings, and postscripts, were applied accordingly.

Along those lines, the qualitative interview began with a short questionnaire that was used to collect demographic information of the interviewee.[37] The short questionnaire was followed by the qualitative interview, and the interview guideline compiled prior to the interview was applied.

Oriented alongside the aim of collecting information on the educational needs of street vendors in Cape Town/South Africa, the interview guideline consisted of three stages, addressing the past, present and future of individual education and work biographies. This structure was chosen because it allowed the interviewee to be introduced to the narration of the topic step by step.[38]

The interview guideline of this study is included in the appendix (tab. 11). It comprise of the following content:

I: Past Work Situation (Motives for working as a Street Vendor)

Focused on the educational and occupational status of the past, questions at level I

[37] The short questionnaire is available in the extended appendix (sec. 'Short Questionnaire').

[38] Interviewee IL05_1 narrated along the structure of the guideline without using it, and thus, reconstructed the guideline intuitively. It can therefore be assumed that the construction of the guideline was meaningful.

specifically asked for information describing the period prior to working as a street vendor selling TBI, and focused on exploring reasons for taking on informal employment. Despite collecting important background information regarding the interviewee's educational and occupational past, level I aimed at preparing the interviewee for the following levels.

II: Present Work Situation (Knowledge and Skills; Seizing Learning Opportunities)
Aimed at exploring the current occupational situation, level II is divided into three sub-levels. First, the interviewee was asked to reflect on the current work situation, and thus, talk about the advantages and disadvantages of the current work status as a street vendor selling TBI. Second, the interviewee was asked to enumerate the acquired qualifications and skills needed for working as a street vendor, and to report on how and where those were acquired. Lastly, the interviewee was asked to evaluate trainings and workshops offered by TBI, as well as individual advantages taken of the participation in respective learning opportunities, especially with regard to moving from the informal to the formal labor market.

III: Future Work Situation (Educational Investments)
Finally, level III addressed the future educational and occupational status. Level III focused the potential outcomes of educational investments, for example participating in learning opportunities offered by The Big Issue. At this point, the interviewees were asked to report on specific wishes and goals concerning their future work situation. While level II addressed necessary skills for working as a street vendor, level III focused educational needs, eventually required to realize such wishes and goals.

The interview ended with an open category that offered the opportunity to make additional comments regarding the past, present, and future educational and work status of the interviewee.

The interviews were recorded with an audiotape recorder and supplemented by a postscript, prepared by the interviewer. For the purpose of analyzing the interview material, all au-

diotape recordings were transcribed.[39]

7.1.4. Limitations of the Study

The aforementioned scarcity of resources to conduct the qualitative study led to some noteworthy limitations, especially affecting sampling procedures and, consequently, the contextual richness of the data material. As stated before, this study employed a homogeneous sample aimed at purposefully sampling individuals from the group of street vendors selling The Big Issue (TBI) in South Africa. However, the convenient sampling of certain characteristics was necessarily ensured as well in order to realize this study.
One of those characteristic was language proficiency. However, although mastering English was a prerequisite to participate in the study in order to communicate with the interviewer during the interview, transcripts showed that 11 out of 12 interviewees had difficulties understanding or answering certain questions of the qualitative interview. This limited the interviews regarding richness of content. Also associated with the necessity for language proficiency is the issue of the sample's representativity with regard to the focus group. While this study sought to employ a purposeful sampling approach conveniently gathering individuals from the socioeconomically homogeneous group of street vendors selling TBI, interviewees who could not speak English well enough were excluded, meaning that the group of street vendors selling TBI is not fully represented by the sample of interviewees in this study.
Another characteristic was the interviewees' belonging to a specific group of the informal sector, namely street vendors selling TBI in Cape Town/South Africa. This is because TBI enabled and facilitated the interviews, which would have not been realizable otherwise. However, by drawing the sample from the group of street vendors affiliated with TBI only, other groups of the informal sector were excluded from the study. This is important to mention because street vending is not considered the biggest sector of informal work

[39] The transcription key is available in the appendix (sec. A.2). Complete interview transcripts and interview postscripts are available in the extended appendix (sec. 'Interview Transcripts'; 'Interview Postscripts').

in South Africa (sec. 2.2.3), and people associated with other areas of the informal sector might have a different perspective on the topic of educational needs, and show different educational investment behavior.
To that end, data of the qualitative study can only provide insights to a specific group of the informal sector. Respective limitations have to be considered for the interpretation of results.

7.2. Methods

7.2.1. Empirically Grounded Construction of Types and Typologies

Types and typologies are an immanent part of qualitative research practices, though their construction has mostly been nontransparent and unsystematic (Kluge, 2000). Aimed at counteracting thereof, Kelle & Kluge (1999, 2000) developed a methodology on the empirically grounded construction of types and typologies in qualitative social research (Kluge, 1999). Respective methodology provides the frame for analyzing qualitative data in this study. The following section briefly emphasizes the premises of the empirically grounded construction of types and typologies, and elaborates on its application with regard to the research questions of this study.

Type, Attribute Space, and Typology

Types and typologies serve the purpose of comprehending, explaining and understanding complex subjects (Kluge, 2000).
To that end, several different types are identified, namely ideal, real, average, and extreme types as well as prototypes. Though, regardless of their characteristics, every type results from grouping one or more attributes. The so-called *combination of attributes* follows two prerequisites. First, each type needs to consist of attributes that complement each other, and thus, lead to internal homogeneity. Types may be constructed along the lines of attitudes unifying individuals or groups, events, situations, and actions (Kluge, 1999). Second, several types need to be substantially different from each other, leading to external homogeneity (Kluge, 2000).

According to Kelle & Kluge (1999, 2010), a type is considered equivalent to a category or an attribute, and the dimensions of attributes [Merkmalsausprägungen] are considered as subcategories. The sum of subcategories is the attribute space [Merkmalsraum]. The identification of an attribute space or the sum of dimensions is referred to as the process of dimensionalization (Kelle & Kluge, 2010).
Finally, a typology is build upon a set of individual types, externally heterogeneous to each other. In order to combine types to a typology, their attributes must be comparable, because "where a basis for comparison is missing, individual types may exist, but no typology" (Lamnek & Krell, 2016, p. 232). Hence, typologies serve as a frame for a set of commonalities and characteristics of a single phenomenon, as well as the diversity of an area that respective phenomena belong to (Kluge, 1999, p. 44).
However, referring to types and typologies as a simple (re-) construction of attribute spaces as well as the combination thereof is an insufficient description of the method (Kelle & Kluge, 1999, p. 90). Rather, the focus is on a purposeful combination of attributes in order to construct empirically grounded types and typologies (Kelle & Kluge, 1999).

The Model of Empirically Grounded Type Construction

The *Model of Empirically Grounded Type Construction* follows a sequence of four stages (Kluge, 2000).

The first stage is the development of relevant analyzing dimensions. This serves the purpose of uncovering similarities and differences between the elements of a type (e.g., individuals, groups, behaviors, actions, events), and thus, to develop meaningful dimensions of attributes for characterizing individual types. The second stage focuses the grouping of cases, and the analysis of empirical regularities. Using the analyzing dimensions developed at stage one, the elements of a type are now grouped. Here, the premises of internal homogeneity of types and external heterogeneity of typologies have to be considered. Stage three focuses the analysis of meaningful relationships between attributes and the construction of types. As a result, attribute spaces are often reduced to few distinguished

types. Finally, the fourth stage is the characterization of the constructed types. This step serves to characterize the individual types on the basis of the combinations of attributes and their relationships (Kluge, 2000).

Lastly, the Model of Empirically Grounded Type Construction is not subjected to linear causality. Especially the formation of multidimensional types and typologies demands a repetition of the aforementioned stages (Kelle & Kluge, 1999). Hence, figure 8 emphasizes the recursiveness of the process.[40]

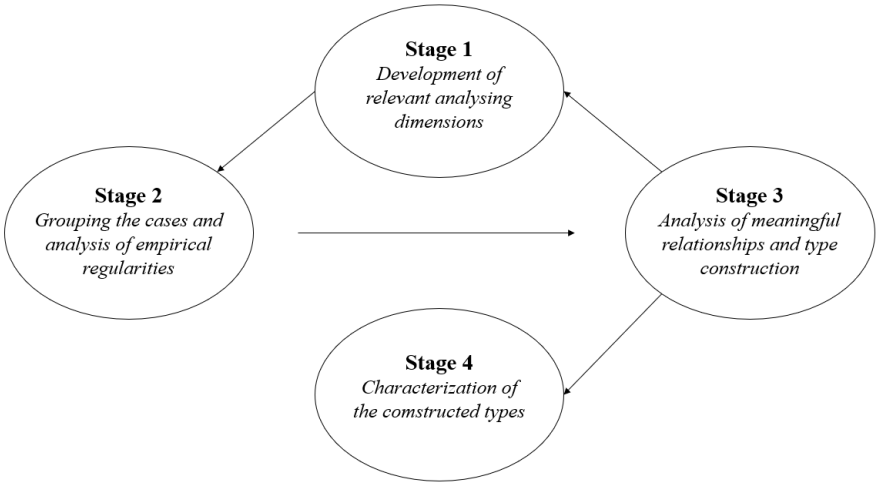

Figure 8: Association of stages of the Model of Empirically Grounded Type Construction after (Kluge, 2000). Adapted figure.

7.2.2. Rational for Method Selection

In order to explore the educational investment behavior of South Africa's socioeconomic low-performers, the Model of Empirically Grounded Type Construction after Kelle & Kluge (1999, 2000) is particularly suitable for various reasons.

First, the flexibility of the method proves to be advantageous. For instance, since this

[40] Kelle & Kluge (2000) point towards different techniques that may be applied in order to realize certain stages of type construction (Kluge, 2000). Respective techniques are summarized in figure 17 in the appendix.

study utilizes secondary qualitative analysis, the relevant analyzing dimensions cannot be derived directly from the guidelines of the qualitative interview. Though, by permitting a variety of analyzing methods and techniques as well as the recursiveness of the stages included in the Model of Empirically Grounded Type Construction, the method enables sufficient data analysis nonetheless. Another example of its flexibility is the method's premise to form dimensions of attributes both, ex-ante/deductively, and ad-hoc/inductively from the data (Kelle & Kluge, 1999). This proves to be particularly beneficial for qualitative secondary analysis, since it is assumed that the interview material collected in the primary study contains a fair amount of information that is irrelevant to the research question of the secondary analysis. An attribute space developed prior to data analysis therefore enables the selection of essential aspects important to the topics of interest, and hence, prevents possible data overload (Miles & Huberman, 1994).
Furthermore, empirically founded typification is insensitive to the number of cases, that is, even a single case can be sufficient for carrying out successful typification (Kelle & Kluge, 1999).
In addition to its flexibility, the Model of Empirically Grounded Type Construction also offers the essential advantage of enabling two levels of analysis. On the one hand, typification functions on a descriptive level, hence it serves to group cases, and thus, contributes to a better comprehensibility of the data. On the other hand, it serves to elaborate on similarities and commonalities in the data, and thus, stimulates the uncovering of mechanisms and contexts within a typology (Lamnek & Krell, 2016).

7.3. Results

In the following section, the qualitative data is analyzed using the Model of Empirically Grounded Type Construction after Kelle & Kluge (1999, 2010), and with regard to the motivational and volitional phase of action behavior as appointed in the Rubicon Model of Action Phases after Heckhausen & Gollwitzer (1987). The qualitative analysis is aimed at addressing research questions I and II of this study (sec. 5). Procedures for analyzing the data material are in line with the different techniques of type construction (fig. 17 in the appendix), though only the procedures that are meaningful and relevant for the analysis are utilized. Furthermore, for the ease of comprehension and along the lines of the aforementioned terminology, types are interchangeably referred to as categories, and dimensions of attributes are referred to as subcategories. In terms of software tools, MAXQDA 10 was used for stages one and two of the Model of Empirically Grounded Type Construction.

7.3.1. Types of Educational Investments

Stage 1: Development of relevant Analyzing Dimensions

For the development of relevant analyzing dimensions to categorize the available interview material ex-ante, Kelle & Kluge (1999, 2000) appoint three procedures, derived from three different types of prior theoretical knowledge (Kelle & Kluge, 2010). Respective procedures are:

1. Empirically unsubstantiated, abstract theoretical concepts
2. Everyday concepts
3. Empirically substantiated, rich concepts

The development of relevant analyzing dimensions on the basis of empirically unsubstantiated, abstract theoretical concepts or everyday concepts with no empirical content is rejected due to the preconditions of the data. For example, everyday concepts provide

a basis for understanding everyday life and usually serve as a starting point for the construction of interview guidelines. Vice versa, they are likely to cause the development of relevant analyzing dimensions along the lines of the categories of the interview guidelines (Witzel, 2000). Since the interview guidelines of this study were developed to suit the needs of the primary study, namely the exploration of educational needs, they can only serve as the basis for the development of analyzing dimensions fitting the topic of the secondary analysis to a limited extend. Although the elements 'Future' or 'Future work situation' and 'Necessary measures for support in the field of education', as included in the interview guidelines, are also relevant for secondary analysis, these topics do not offer a framework sufficiently flexible for the exploration of educational investment behavior in the focus group.

Accordingly, the analyzing dimensions or categories used in this study are formed on the basis of empirically substantiated, rich concepts. A concept is considered empirically rich if it clearly differs from other concepts, and thus, acquires empirical validity. Empirically rich concepts are often derived from other subject areas and studies, and can be confirmed or refuted in the course of analysis (Kelle & Kluge, 2010). Hence, observations from exploring educational needs are used accordingly to develop empirically meaningful concepts useful for secondary analysis of the data.

Of particular importance is the dilemma *earning vs. learning*. Results of the primary study on educational needs showed that interviewees cannot afford to invest time or monetary resources in educational attainment. More specifically, educational attainment would entail that street vendors loose time to sell The Big Issue, and hence, forgo valuable income (Twele (formerly Radermacher), 2013). Although, the motivation for educational investments is given, it is not yet transferred into actions due to the dilemma of having to earn money for a living, and wanting to invest in education. Types of educational investment behavior of the focus group are therefore based on two attributes. Those are the motivation to invest in education, and the ability to transfer respective motivation into actions. To capture thereof, two categories are formed, namely *Planned Investments* and *Placed Investments*. The distinction between the two also aligns with the phases of

motivation and volition, manifested in the theoretical framework of this study (sec. 3).

For the dimensionalization of the categories Planned Investments and Placed Investments, empirically founded subcategories are constructed, based on the interview material of the original study. This step serves the purpose to substantiate an enrich the analyzing dimensions (Kelle & Kluge, 2010). When defining the subcategories, care must be taken to establish a clear demarcation of similarities and differences between the individual cases, in order to ensure internal homogeneity and external heterogeneity.
Because this study focuses educational investment behavior of a specific group, dimensionalization is performed across cases. Furthermore, dimensionalization is performed ad-hoc/inductively, and therefore results from the data itself. Figure 9 summarizes categories as well as corresponding subcategories, namely analyzing dimensions. Values in square brackets indicate the number of statements counted for specific categories and subcategories.

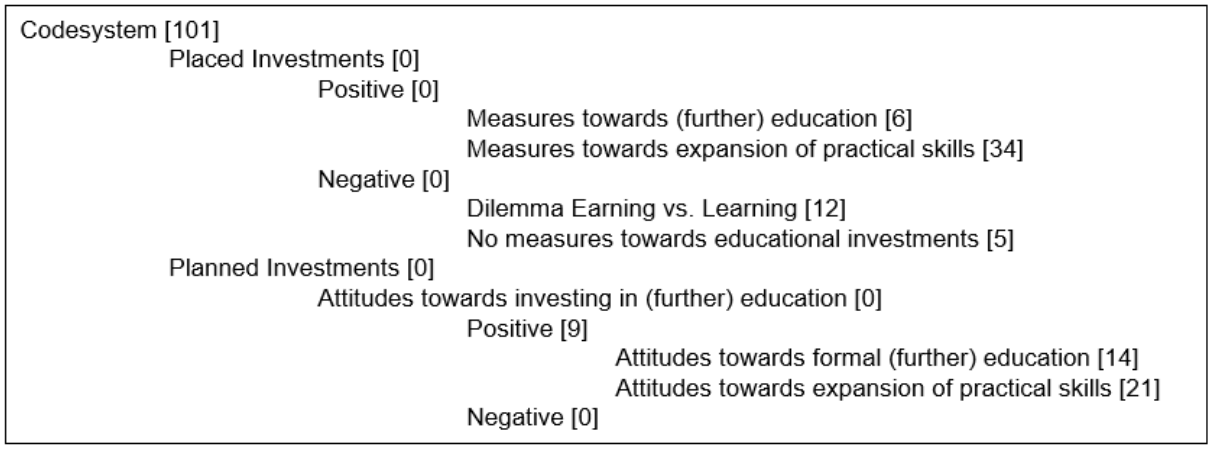

Figure 9: Analyzing dimensions on the topic of educational investment behavior. Figure obtained from MAXQDA 10, based on own analysis.

The first category (Placed Investments) addresses actional investment behavior, and is divided into two subcategories based on the data material, namely positive and negative investments. The former refers to investments that have been taken, and the latter

addresses those that have not been taken. That is, the *negative* investments refer to non-investments. This subdivision is in line with the logic of comparing a phenomenon through categories, entailing at least two alternating options (Kelle & Kluge, 2010), and thus, corresponds to the premises of EV-Theory after Esser (1999), stating that each action decision is based on at least two action alternatives (sec. 3.2.2).

The subcategories of positive and negative investments are further specified. Based on the data, positive Placed Investments are subdivided into *measures towards (further) education*, summarizing actions that relate to continuing formal education of any sort, including catching up on obtaining a diploma or signing up for college; and *measures towards expansion of practical skills*, summarizing actions that foster the acquirement or the expansion of practical skills, for instance through the participation in workshops, or by learning new skills. Furthermore, negative Placed Investments are subdivided into the *dilemma earning vs. learning*, accounting for the dilemma of being motivated to take educational investments, but being unable to do so due to a shortage of time and monetary resources; and *no measures towards educational investments*, addressing educational non-investments due to no particular reason.

The second category (Planned Investments) focuses attitudes towards investing in (further) education, and is divided into positive and negative attitudes. Again, this is to account for complimentary pairs of action alternatives. The subcategory of positive attitudes is subdivided into *attitudes towards formal (further) education* and *attitudes towards expansion of practical skills*. With regard to the subcategory of negative attitudes, the data did not show cases of interviewees indicating to be unmotivated to invest in (further) education. Hence, the subcategory is found empty, and thus, dropped from the attribute dimension accordingly.

Figure 10 shows the results of the MAXQDA Code-Matrix-Browser (CMB), which was used to calculate the frequency of statements from interviewees assigned to codes (i.e., categories and subcategories). Squares indicate the number of statements related to each code (printed in the rows) occurring in each case (printed in the columns). Bigger squares

indicate a larger number of statements related to a code, and smaller squares indicate the opposite. In this study, the matrix was computed based on the rows (MAXQDA, 2011).[41]

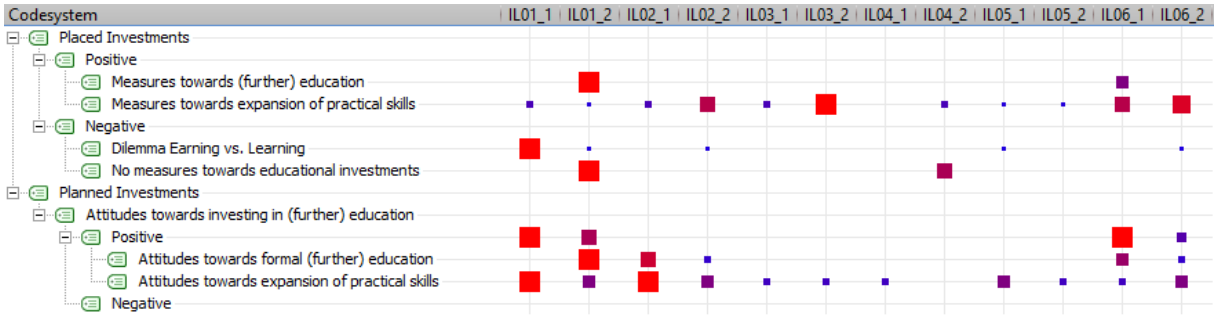

Figure 10: Distribution of cases across analyzing dimensions. Figure obtained from MAXQDA 10, based on own analysis.

Stage 2: Grouping the Cases and Analysis of Empirical Regularities

Given that types of educational investment behavior entail the combination of subcategories on motivational attitudes and volitional behavior towards educational attainment, a multidimensional typology is indicated. In this case, Kelle & Kluge (1999, 2010) argue for the application of an attribute space, typically realized through a cross table. The latter serves the purpose of a systematic and transparent sorting of existing data and the empirically driven grouping of cases (Kelle & Kluge, 2010). Figure 11 shows the attribute space of this study, that is, a cross table summarizing the developed categories and subcategories, providing an overview of all theoretically possible and empirically derived combinations between Planned Investments and Placed Investments.

The average count of cases per combination of attributes equals $\bar{x} = 3.38$, excluding negative Planned Investments given that none of the interviewees referred to that category. The attribute space shows that the association between Planned Investments, namely

[41] Note that the matrix displays the personal code of each interviewee in the columns, for example, IL01_1. For detailed information on the personal code see the transcription key in the appendix (sec. A.2).

Placed Investments	Planned Investments		
	Attitudes towards investing in (further) education		
	Positive		Negative
	Attitudes towards formal (further) education	Attitudes towards expansion of practical skills	-
Positive			
Measures towards (further) education	2	2	-
Measures towards expansion of practical skills	5	9	-
Negative			
Dilemma Earning vs. Learning	2	5	-
No measures towards educational investments	1	1	-

Figure 11: Distribution of cases across the combination of attributes in the attribute space. Own illustration, based on own analysis.

positive attitudes towards expansion of practical skills, and Placed Investments, namely positive measures towards expansion of practical skills is observed more frequently than all other combinations (n = 9). Furthermore, the association between the subcategories attitudes towards formal (further) education and measures towards expansion of practical skills (n = 5), as well as attitudes towards expansion of practical skills and dilemma earning vs. learning (n = 5) is above average. On the other hand, the association between the subcategories of positive attitudes towards formal (further) education and no measures towards educational investments, is below average indicating a less frequent combination of these subcategories. Thus, results imply that positive attitudes for educational investments are more often followed by placed investments.

Overall, it is found that positive attitudes towards investing in (further) education result more frequently in educational investments than in educational non-investments. Against the background of the focus group's characteristics as shown in section 2, and the theory-driven assumption that South Africa's socioeconomic low-performers are more likely to

Placed Investments	Planned Investments	
	Attitudes towards investing in (further) education	
	Positive	
	Attitudes towards formal (further) education	Attitudes towards expansion of practical skills
Positive		
Measures towards (further) education	Type I *Targeted*	
Measures towards expansion of practical skills		
Negative		
Dilemma Earning vs. Learning	Type II *Thwarted*	
No measures towards educational investments	Type III *Inconsistent*	

Figure 12: Empirically grounded construction of types on educational investment behavior among socioeconomic low-performers in South Africa. Own illustration, based on own analysis.

refrain from taking educational investments due to the foundations of cost-benefit calculations (sec. 3.2.2), the results of the attribute space do not correspond to theoretical assumptions along the lines of rational decision making. Respective results are discussed below.

Stage 3: Analysis of meaningful Relationships and Type Construction

Results obtained from the attribute space are now summarized into types. The aim of this stage is to discover social structures that are represented by the combination of subcategories or attribute spaces respectively. At this point, the latter is typically reduced, and observed combinations of subcategories are summarized to distinctive types (Kelle & Kluge, 2010). Figure 12 presents the final three types of educational investment behavior, based on the results of the attribute space. Those are the types which can be described as (I) *Targeted*, (II) *Thwarted*, and (III) *Inconsistent*.

Type I *Targeted* represents target-oriented educational investment behavior. Interviewees

acting according to type I are motivated to take educational investments, and purposefully invest in both formal and informal education. This is achieved by participating in workshops offered by The Big Issue, as well as by participating in the formal education system and, for example, signing up for college to obtain a diploma.
Type II *Thwarted* represents educational investment behavior that has been obstructed by external factors. Interviewees acting according to type II have a positive attitude towards investing in (further) education. That is, they are motivated to take educational investments. However, external factors such as the dilemma earning vs. learning prevent the transformation of Planned Investments into Placed Investments, and thus, the transformation of motivation into volitional behavior.
Type III *Inconsistent* represents inconsistent educational investment behavior. Like type I and type II, interviewees acting according to type III show positive attitudes towards investing in (further) education. However they refrain from acting on a motivation and do not take educational investments. Type III is different from type II since interviewees acting according to type III do not state reasons why motivation is not transferred into action.

Throughout the analysis, it became evident that all three types identified here are grounded on positive attitudes towards investing in (further) education, that is the motivation for educational investments. Thus, the premise of external homogeneity is first redeemed when respective motivation is linked to action, that is under consideration of Placed Investments.

Stage 4: Characterization of the Constructed Types

Lastly, types of educational investment behavior are characterized in more detail. Focus is put on the premise of internal homogeneity, as well as the context of meaning with regard to each type. To that end, the most concise statements obtained from the qualitative interviews are referenced here.[42]

[42] If part of the statement has consciously been left out in this section, [...] has been used to indicate that. A complete overview of the most comprehensive statements is available in table 12 in the

Type I *Targeted* represents active, conscious and goal-oriented educational investment behavior. Respective type consists of positive attitudes towards investing in (further) education, that is positive attitudes towards formal (further) education and the expansion of practical skills. Compared to the other types, type I shows the most active investment behavior, given that interviewees acting according to type I not only express the necessity for educational investments, but also continue to take them. Interviewees acting according to type I invest their resources in different areas of educational attainment. For example, by participating in the informal education program offered by The Big Issue, including sales and business training but also workshops on child- and healthcare among others; as well as investments in formal education, namely pursuing (further) educational training and aiming at obtaining a diploma. For example, interviewee IL06_1 reported that he enrolled to go back to school and complete *Further Education and Training* (sec. 2.2.4):

> "Yes so I'm gonna do grade 12 there at Philippi next year." (IL06_1)

Along the same lines are investments taken by IL01_2, who was already enrolled in college:

> "Ehhm to put me further up my study. I was in technical college of [-]. [...] I was being industrial electronics electric theory ehm math and all other subjects." (IL01_2)

Lastly, statements of interviewees IL01_1 and IL02_2 are used as examples for measures taken towards the expansion of practical skills, such as the participation in workshops offered by TBI.

> "Yes I tried to be participating in some workshops. I'm doing mosaic art." (IL01_1)

> "I attend all the workshops." (IL02_2)

---

appendix and a summary of all statements assigned to a category or subcategory is available in the extended appendix (sec. 'Summary of Statements').

On the contrary, interviewees acting according to type II *Thwarted* and type III *Inconsistent* do not take educational investments, although they acknowledge the necessity thereof and are motivated to do so. Along those lines, type II represents an educational investment behavior that is obstructed by external factors, such as the dilemma earning vs. learning. On that note, interviewee IL01_1 emphasized the dilemma as follows:

> "But I believe we need [-] allowances we need allowances [-] so that you won't you won't be stuck. You won't starve by at least doing those learnerships. [...] Some people they don't have the time. They want to be working they want to be learning. [...] But now instead of working and learning. Which one would you choose? They obviously choose working where they'll be earning." (IL01_1)

Lastly, type III *Inconsistent* represents an inconsistent educational investment behavior. Interviewees acting according to type III have recognized the necessity for educational investments, thus, in contrast to types I and II, corresponding investments are not taken. A potential reason for non-investments is the lack of relevant learning opportunities as expressed by Interviewee IL1_2:

> "So I don't want to lie there have been so many offers but I haven't attended. There hasn't been the one that I will specifically like. [...] They do offer us like computer courses and all that. I don't want a computer course." (IL01_2)

7.3.2. Summary

To summarize, findings of secondary qualitative analysis resulted in three types of educational investment behavior, namely type I *Targeted*, type II *Thwarted*, and type III *Inconsistent*.

Contradicting the theoretical assumptions outlined in section 3, all three types showed positive attitudes towards educational investments and are hence considered motivated.

Though, investment behavior according to type II and type III does not entail the transformation of respective motivation into volition, namely educational investments. In the case of type II, primary reasons for non-investment are the dilemma earning vs. learning or the lack of time and monetary resources. This is in line with the argumentation provided by Becker (1962) among others, stating that the element of time is particularity decisive for making rational decisions on an educational investment (sec. 4).
The most relevant result for the scope of this study is investment behavior according to type I *Targeted*. This is because interviewees acting along the lines of type I are not only motivated to invest in education but also take educational investments. That is, regardless of their socioeconomic status or the availability of time and monetary resources. For example, actions of educational investments according to type I encompass the participation in the education program offered by The Big Issue, mainly focused on building life skills such as knowledge in sales and business as well as childcare and healthcare, and investments in formal education, namely pursuing further schooling.
To summarize, results of qualitative analysis are in favor of hypothesis $H_{1.1}$ (sec. 5). Furthermore, type I *Targeted* also supports $H_{2.1}$. However, given that not all street vendors in the sample turn Planned Investments into Placed Investments, and hence, motivation into action, the null hypothesis $H_{1.0}$ is true for type II *Thwarted* and type III *Inconsistent*.

As previously mentioned, findings of the qualitative study are not considered representative. They merely provide an insight to patterns of educational investment behavior among street vendors participating in the interviews of the qualitative study. Regardless thereof, types of investment behavior deducted from the qualitative data interestingly showed that all interviewees have positive attitudes towards educational investments, and are considered highly motivated. In the case of type I, interviewees even transform respective motivation into actions.
On that note, findings of the qualitative study revealed the need for further analyses that distinctively look at separate phases of educational investments behavior, identify their potential effects, and allow for generalization of assumptions on the educational

investment behavior of socioeconomic low-performers in South Africa. To that end, the following quantitative study is conducted.

8. Quantitative Study: Parental Educational Investments

Subsequent to the qualitative perspective, this section presents the quantitative perspective of this study, emphasizing the effects of parental involvement activities as a form of educational investment behavior.

The theoretical framework for the quantitative study is the Model of Parental Involvement Process, aligned with the Rubicon Model of Action Phases (sec. 3). Whereas the qualitative study focused the first two phases of action behavior with regard to overall educational investments among street vendors in Cape Town/South Africa, the quantitative study reflects all three phases of the model, and addresses motivation, volition, and the evaluation of actions.

In the first part of this section, the data set utilized for the quantitative study is briefly described, accompanied by references to official documents providing in-depth information thereon (e.g., references provided by Boston College and the IEA).[43] The second part of this section gives an overview of the empirical methods applied in this study, and special attention is paid to issues particularly relevant to the conducted analysis. This includes comparing different population groups, and addressing large shares of missing information occurring in the utilized data set. Lastly, descriptive information are reported, and results of the quantitative study are presented.[44]

[43] Data for the quantitative analysis is the merged and recoded South African data of the international data base of prePIRLS 2011 (PIRLS and prePIRLS 2011 Assessment. Copyright ©2013 International Association for the Evaluation of Educational Achievement (IEA). Publisher: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College, available at: https://timssandpirls.bc.edu/pirls2011/international-database.html).

[44] Descriptive statistics presented in this section are computed with IEA's International Database Analyzer (IDB Analyzer, version 4.0.39). The IDB Analyzer is a software tool compatible with SPSS and build to combine and analyze data from large-scale assessments in an appropriate manner, taking into account "sampling information and the multiple imputed achievement scores to produce accurate statistical results" (Foy & Drucker, 2013b, p. 5). Statistics concerning factor analysis, invariance testing, and structural equation modeling are computed with the software Mplus (version 8.01).

8.1. Data

For this study, South African data of prePIRLS 2011 collected via procedures of the *Progress in Reading Literacy Study* (PIRLS) is used.

Even though the more recent data set of PIRLS 2016 is already available, the 2011 prePIRLS data set has consciously been chosen due to three reasons. First, the qualitative intake of this study is based on data collected in 2013, and thus, the timely proximity of both, the quantitative and the qualitative study, is ensured. Second, South African fourth-grade students who participated in prePIRLS 2011 took the test in one of the 11 official languages acknowledged by the South African government, allowing for this study to draw on a large sample of n = 15,744 students (Howie et al., 2017). This is particularly beneficial to the applied method of multi-group structural equation modeling on non-normally distributed, categorical data using maximum likelihood estimation (e.g., Bandalos, 2014; Savalei & Bentler, 2005). Third, associations between parental involvement and student achievement are found to be stronger for younger students (Caro, 2011), and hence, effects of parental involvement might be more perspicuous for fourth-grade students. Respective reasons also justify the use of prePIRLS over PIRLS 2011.

8.1.1. Features and Limitations of PIRLS and prePIRLS

PIRLS is an international comparative cross-sectional study carried out by the IEA on a five-year schedule. First conducted in 2001, PIRLS focuses the assessment of student reading literacy achievement at grade 4 on an internationally comparable scale. The study consists of a student assessment on reading as well as several background questionnaires for students, parents, teachers, and principals collecting contextual information on the educational environment of students. Furthermore, representatives and governing entities inform about country-specific educational systems and reading curricula (I. V. S. Mullis et al., 2009).

All assessments and background questionnaires are provided in an international version, which is translated by "participating countries into their languages of instruction with the goal of creating high quality translations that are appropriately adapted for the national

context and at the same time are internationally comparable" (I. V. S. Mullis, Drucker, et al., 2012, p. 5).

By design, PIRLS aims at assessing the proficiency of students at points of transition, that is moving from primary to secondary education (sec. 3.2.2). Aimed at ensuring comparability between the performance of the target population across education systems participating in PIRLS, UNESCO's International Standard Classification of Education (ISCED) is applied to sample students that have received the same level of education across countries (Joncas & Foy, 2012). ISCED is part of the *United Nations International Family of Economic and Social Classifications*, and serves the purpose of "assembling, compiling and analyzing cross-nationally comparable data" (UNESCO, 2012, p. 6). Since 2011, the classification system consists of eight levels, covering all stages of formal schooling, that is from early childhood education equaling level 0, to the second level of tertiary education equaling level 5 and above (UNESCO, 2012). The classification was revised in 2011, however PIRLS 2011 has been built upon the revisions of 1997. The PIRLS assessment targets students of ISCED level 1. In most countries, that is the case for students in grade 4, aged about 9.5 years (I. V. S. Mullis et al., 2009). However, in some education systems, students collectively reach the expected level of reading proficiency only at a later stage, for example at grade 5 or 6. This is often the case in developing countries. To accommodate this particularity, and still enable respective countries to assess their students in reading and literacy performance, the IEA introduced a less difficult assessment, namely prePIRLS 2011. The study is based on the same framework of reading comprehension as the PIRLS assessment, and thus, paves the way towards participating in the main study. Given that prePIRLS was designed in parallel to PIRLS, the study made use of the same context questionnaires. Interestingly, prePIRLS 2011 was piloted in South Africa (I. V. S. Mullis, Drucker, et al., 2012).

In total, the PIRLS 2011 data set holds information on approximately 325,000 students from 49 countries and nine benchmark systems (I. V. S. Mullis, Martin, Foy, & Drucker,

2012).[45] In South Africa, a sample of approximately 20,000 students in over 400 schools was selected (sec. 8.1.2). In 2011, both fourth-grade and fifth-grade students were assessed with regard to their reading proficiency. Fifth-grade students took the PIRLS assessment in either English or Afrikaans. Thus, this grade level functioned as a benchmark to collect information on language of instruction, namely English or Afrikaans (I. V. S. Mullis, Martin, Foy, & Drucker, 2012). On the other hand, fourth-grade students took the less difficult prePIRLS assessment in the language of instruction that they experienced up until grade 3, namely one of the 11 official languages acknowledged by the South African government (Howie et al., 2012). This resulted in a total of 15,744 fourth-grade students in the South African 2011 prePIRLS data set.

Although advantageous in many ways, for example regarding sample size and richness of information, large-scale assessments such as PIRLS and prePIRLS come with some noteworthy limitations that need to be considered when analyzing the data and interpreting the results.

For instance, both PIRLS and prePIRLS are cross-sectional studies. Trends are only available through aggregated data at country level. Consequently, analyses based on PIRLS and prePIRLS data do not allow for any causal inferences. Although referred to as effects, associations analyzed in this study are merely non-directed correlations. Furthermore, predefined information collected in standardized assessments such as prePIRLS limit the operationalization of theoretical frameworks, and often call for the usage of approximations to the original construct. The latter is true for this study, and hence, some adjustments to the original theoretical constructs have to accepted (sec. 8.2.3).

[45] Benchmark systems are referred to as sub-national administrative entities. For example, Canada participated as one educational system, as well as with the Canadian provinces - namely Alberta, Ontario, and Quebec - as benchmark systems (Rutkowski et al., 2014).

8.1.2. Sampling Procedure and Target Population

In general, PIRLS ensures a standard of national representative samples including at least around 4,000 students from 150 - 200 schools (I. V. S. Mullis, Martin, Foy, & Drucker, 2012). To achieve that, PIRLS employs a multi-matrix sampling technique (I. V. S. Mullis et al., 2009), namely a stratified two-stage random sample design (Joncas & Foy, 2012). The initial units for the process of stratification are schools within an educational system. At stage one of the sampling process, schools are sampled according to the premises of explicit and implicit stratification. In the process of explicit stratification, schools are grouped by characteristics such as regions or provinces of an educational system, resulting in several explicit strata. Furthermore, implicit stratification is employed to sort schools *within* explicit strata, for example, according to school type (Joncas & Foy, 2012). Explicit and implicit stratification at stage one of the sampling process is then followed by a selection of classes and students within schools at stage two. Respective selections are conducted according to the premises of systematic random sampling (Joncas & Foy, 2012). Assuming that students' curricular and instructional experiences are more sufficiently captured within rather than across classrooms of one grade, students nested within one classroom are sampled. Additionally, participating countries need to define a student target population in order to ensure a national representative sample of schools, classes, and students. Sampling frames might deviate internationally due to national specifics with regard to the education system and curriculum. Furthermore, as a consequence of collecting a two-stage random sample, PIRLS data is hierarchically structured, and not all students are given the same probability to be part of the sample. Therefore, PIRLS data is provided with sampling weights adjusting for school and student samples with regard to the total population coverage (Joncas & Foy, 2012).

For prePIRLS 2011, South Africa opted to oversample the population of fourth-grade students, aiming to achieve more precise estimates for each of the 11 languages of instruction (Martin & Mullis, 2012a). To that end, students were assessed in the language they have been instructed in during the first three years of schooling (van Staden et

al., 2016). Explicit stratification was based on grade levels and language of instruction. Implicit stratification was based on language profiles of schools (i.e., one language, two languages, and multi-language schools), as well as South African provinces (Martin & Mullis, 2012a). South Africa sampled one classroom per language per school, meaning two classrooms were sampled in bilingual schools. Furthermore, PIRLS and prePIRLS were administered in the same schools, ensuring an overlap between the samples. In total, 352 schools participated in prePIRLS and PIRLS 2011, 345 of which administered prePIRLS (Martin & Mullis, 2012a).

As a result, prePIRLS achieved to cover 97% of the target population, which are South African fourth-grade students, with only 2.1% excluded at school level. Exclusions only concern very small schools as well as educational institutions declared as language schools, focusing on languages of instruction different from the 11 official South African languages. The within-school exclusion rate is below 1%, and thus, considered negligible (Joncas, 2012; Martin & Mullis, 2012a). Generally, exclusion rates in South Africa's prePIRLS sample of 2011 are noncritical given that they remain below the 5% threshold (Joncas, 2012). On that note, the South African sample of prePIRLS provides a solid base for empirical analyses reflecting on the target population.

8.1.3. Plausible Values

Aimed at comprehensively measuring reading literacy within the target group, PIRLS and prePIRLS draw back on an extensive pool of literary and informational reading passages, as well as corresponding questions of comprehension. However, the entirety of the PIRLS and prePIRLS assessment would exceed the cognitive abilities of a single student, risking loss of concentration and fatigue effects (I. V. S. Mullis et al., 2009). To mitigate such risks, and to keep the burden on students completing the assessment within a reasonable time span, both studies rely on a matrix-sampling booklet design. That is, students are given systematically arranged test booklets, consisting of a combination of test blocks, comprising either literary or informational tasks and items. In prePIRLS 2011, 123 items were distributed to a total of six test blocks, rotated among nine test booklets. To safe-

guard the linkage between test booklets, each block appeared in three different booklets. Thereby, the matrix-sampling design advantageously ensures a limitation of the total test time to 80 minutes per student without compromising the comparability or the contextual richness of the test results (I. V. S. Mullis et al., 2009).

In order to obtain estimates of student performance in reading from the assessments, the following scaling methodology is applied.
First, an item response theory (IRT) scaling approach is utilized to ensure an accurate description of student achievement on the test as well as measures of trend. In total, three IRT models are applied, accounting for the item characteristics in PIRLS and prePIRLS, and building on the association between student proficiency and the probability that a student will respond to an item in a certain way (Foy et al., 2012).
Second, multiple imputation, also known as plausible values, is applied in combination with conditioning to obtain proficiency scores in reading for all students participating in the PIRLS and prePIRLS assessment. The premise of this approach is to use all available data, namely student responses as well as available background information, in order to estimate student ability. Ideally, this approach also levels a potential measurement error occurring due to the matrix-sampling booklet design (Foy et al., 2012).

Lastly, the assessments of PIRLS and prePIRLS result in five plausible values, representing the overall reading literacy scale, as well as plausible values reflecting on sub-scales of reading comprehension at ISCED level one (I. V. S. Mullis et al., 2009). However, given the scaling methodology applied in the studies, reading literacy scales represented by plausible values "are not intended to be estimates of individual student scores, but rather are imputed scores for like students - students with similar response patterns and background characteristics in the sampled population - that may be used to estimate population characteristics correctly" (Foy et al., 2012, p. 3). Plausible values are utilized as outcome variables in this study (sec. 8.2.3).

8.2. Methods

The following section provides an overview of quantitative methods applied in this study.[46] First, the premises of structural equation modeling are highlighted, emphasizing estimation methods and fit indices. Second, premises of invariance testing essential to group comparisons are introduced. Third, the selection of variables for the model of this study is listed. Fourth, measures for group identification are presented. And fifth, the issue of missing data is emphasized, given that variables used for this study are affected by partially high rates of missing information.

8.2.1. Structural Equation Modeling

In this study, structural equation models (SEM) build with the software Mplus (version 8.01) are utilized to explore the association between motivation, volition, and the evaluation of actions in the scope of parental involvement as a form of educational investment behavior.

General Notions on Structural Equation Modeling

Structural equation modeling is a method situated within the family of multivariate methods. In reference to Bentler (1988), Byrne (2012) defines SEM as a procedure where causal processes between multiple items are established (Byrne, 2012):

> "The term structural equation modeling conveys two important aspects of the procedure: (a) that the causal processes under study are represented by a series of structural (i.e., regression) equations, and (b) that these structural relations can be modeled pictorially to enable a clearer conceptualization of the theory under study. The hypothesized model can then be tested statistically

[46] Given similarities of the research topic as well as data and methods, this study is guided by the work of Feld (2018), who established a multi-group structural equation model comparing parental involvement in association with social background in Canada and Germany based on PIRLS 2011 data (Feld, 2018).

> in a simultaneous analysis of the entire system of variables." (Byrne, 2012, p. 3)

Compared to other multivariate approaches such as exploratory factor analysis and regression analysis, SEM comes with three notable advantages.

First, it allows for the incorporation of both, unobserved (i.e., latent) and observed (i.e., manifest) variables (Byrne, 2012). Second, SEM provides more accurate estimates of residual variances, which are error variances. Traditional procedures such as regression analysis on the other hand factor out errors in the predicting variables, SEM "provides explicit estimates of these error variance parameters" (Byrne, 2012, p. 3). SEM therefore proactively corrects for the measurement error (Byrne, 2012). Third, SEM engages a confirmatory rather than an exploratory approach, and is therefore better suited for hypotheses testing and inferential analyses (Byrne, 2012). To that end, SEM is an inimitable method for analyzing complex path models, engaging several elaborate statistical features (Byrne, 2012).

Finally, as most empirical methods in social sciences, SEM requires a strong theoretical framework in order to provide robust results. On that note, Bentler & Chou (1987) emphasized the importance of "theoretically appropriate operationalizations of variables" and latent constructs (Bentler & Chou, 1987, p. 82) when the aim is to build a meaningful empirical model and provide sophisticated evaluation of fit and interpretation of results.

Characteristically, SEMs can be separated into two component models, which are the measurement model and the structural model, also referred to as the latent variable model or factor analytic model (Byrne, 2012; Geiser, 2011).

First, the measurement model represents the association of observed items with a latent factor. Generally it is assumed that the covariance between observed items is due to the latent factor. Hence, changes in the latent factor are coherent with changes in the observed items, and thereby explain associations between the latter. Lastly, associations between the factor and the observed items are specified as regression paths. Coefficients of such associations are often referred to as factor loadings. Second, the structural model

looks at the association between measurement models with each other or with additional manifest variables. Respective associations can be specified as directional paths in the form of regressions, and as non-directional paths in the form of covariance or correlations (Geiser, 2011). Particular to SEM is the simultaneous estimation of the measurement and the structural model.

In order to emphasize the association between (latent) dependent and independent variables in a SEM, path diagrams are commonly utilized. Usually, the following notation is applied.
Latent factors are mapped by circles or ellipses, and observed items are mapped by squares. Directional associations are mapped by single-headed arrows, and non-directional associations are mapped by double-headed arrows. Variables that send arrows are predictors. Variables that only function as such are furthermore characterized as exogenous variables. Variables that receive arrows are referred to as dependent or endogenous variables (Geiser, 2011).
Moreover, single-headed arrows are also used to map the occurrence of errors. Byrne (2012) pointed towards the distinction of two types of errors, which are measurement errors and residual errors. Measurement errors are associated with the observed variables, reflecting "their adequacy in measuring the related underlying factors" (Byrne, 2012, p. 11). Residual errors, on the other hand, occur from the association between endogenous and exogenous factors. They reflect "the extent to which this predicted value is in error" (Byrne, 2012, p. 11). It is important to note that Mplus refers to both types of errors as residual errors.

Application of Structural Equation Modeling in Mplus using prePIRLS

Aimed at building a multi-group SEM, several model specifications required for the analysis of large-scale assessments need to be acknowledged.
First, the specific data structure resulting from student outcomes in reading reflected by five plausible values derived from multiple imputation (sec. 8.1.3) needs to be accounted

for. To that end, the Mplus command *TYPE = IMPUTATION* is utilized. In order to account for five plausible values, five separate data sets are needed, each containing one plausible value. The Mplus command is then applied to draw on each data set and carry out five separate analyses. As for the results, "estimates are [then] averaged over the set of analyses, and standard errors are computed using the average of the standard errors over the set of analyses and the between analysis parameter estimate variation" (L. K. Muthén & Muthén, 1998-2017, p. 509).

Second, the command *TYPE = COMPLEX* is applied in order to account for the nested structure of the data, and thus, ensure the correct estimation of the standard errors and model estimates (L. K. Muthén & Muthén, 1998-2017). By doing so, "corrections to the standard errors and Chi-Square test of model fit that take into account stratification, non-independence of observations, and unequal probability of selection" are obtained (L. K. Muthén & Muthén, 1998-2017, p. 20). More precisely, the *Root-Mean-Square Error of Approximation* (RMSEA) is estimated more accurately, and the Chi^2 test statistic is unlikely to be inflated (L. K. Muthén, 2011). The application of *TYPE = COMPLEX* also requires information on stratification, clusters, and weights. For this study, the stratification variable is *JKZONE*, which informs the software about the assignment of students to sampling zones in prePIRLS. It needs to be noted that *JKZONE* is used alternatively as a stratification variable here since information on explicit and implicit strata are not included in the public use files of the data set. Furthermore, the cluster variable *JKREP* identifies whether a certain "case is to be dropped or have its weight doubled for each set of replicate weights" (Foy & Drucker, 2013b, p. 46). In other words, the use of *JKREP* allows for replicate weights to be applied.[47] Weight is *TOTWGT*, that is the total student sampling weight summing to the student population size in a country. *TOTWGT* is designed to be applied in student-level analyses (Foy & Drucker, 2013b). Technically speaking, it is a "combination of weighting components reflecting selection probabilities

[47] For a technical description of *JKZONE* and *JKREP* along the lines of the 'jackknife repeated replication' (JRR) method used for the computation of sampling variance in PIRLS and prePIRLS see section 7 of the inaugural PIRLS 2011 User Guide (Gonzalez & Kennedy, 2003).

and sampling outcomes at three levels - school, class, and student" (Joncas & Foy, 2012, p. 13). The weighting components adjust for the probability of selection at each level as well as non-participation (Joncas & Foy, 2012). Thus, the total student sampling weight is preferably used in the case of oversampling since it provides each case in the sample with the appropriate weight accounting for its share in the population. Recall, South Africa opted for the administration of an oversample in prePIRLS 2011 to account for language groups.

Third, *Maximum Likelihood Estimation* (ML) (operationalized by the robust ML-estimator MLR in Mplus) is applied for the purpose of this study. The following section emphasizes on that.

Estimators

Typically, data collected in human sciences, and particularly social or educational sciences is mostly obtained from categorical information, which are items with nominal and ordinal scale levels. Furthermore, items are often polynomial, with approximately four answering categories, and item data is rarely normally distributed. All of the aforementioned characteristics also apply to the items of prePIRLS 2011, and hence, interfere with the precise estimation of SEMs. For this reason, the following considerations are relevant.

Scientific voices of the past argued that categorical items are problematic due to the lack of a true numeric nature, which is the basis for mathematical operations applied in SEM. However, empirical studies were able to show that it is not the scale level that is important for such mathematical operations, but rather the number of categories per item. Bentler & Chou (1987) as well as Kühnel (1993) argued that at least four categories are needed for reliable estimations (Bentler & Chou, 1987; Kühnel, 1993). However this is only the case if data is normally distributed (Bentler & Chou, 1987). Similar findings were presented by West et al. (1995), who conducted various simulation studies, and reasoned that the skewness of the data is influential to the Chi^2 test statistic; and a number of categories less than four leads to an underestimation of the measurement model, even when data

is normally distributed (West et al., 1995). Furthermore, the authors found that skewed items with less than four categories caused biased estimations of residual variance. Lastly, West et al. (1995) pointed towards the issue of obtaining underestimated standard errors when working with skewed data (West et al., 1995).

Advantageously, research of the past decades as well as modern techniques of model estimation allow for a correction of respective violations caused by the application of categorical and skewed data within the premises of SEM. For instance, a number of categorical items are applied in the model of this study (sec. 8.2.3). To treat respective data appropriately, the *Weighted Least Squares* (WLS) estimator, and particularly the advancement thereof, the *Weighted Least Square Means and Variance adjusted* (WLSMV) estimator, would be a suitable option (Kline, 2016; Scherer, 2020). WLSMV is known to handle large numbers of subjects and factors well (B. O. Muthén et al., 2015), and is furthermore suited to treat categorical and non-normally distributed data (Urban & Mayerl, 2014). However, the WLSMV estimator does not appropriately account for large amounts of missing information (B. O. Muthén et al., 2015). However, since missing data are a prominent issue in this study (sec. 8.2.5), ML estimation, operationalized by the robust ML-estimator MLR in Mplus is applied in the model of this study instead.[48]
Other than WLSMV, MLR estimation in Mplus takes into account the advantages of data being missing at random (sec. 8.2.5), and builds the estimation on information gained from both covariates and outcomes. MLR computes standard errors using a sandwich estimator, which is the robust covariance matrix estimator yielding reliable estimates "even when the fitted parametric model fails to hold or is not even specified" (Kauermann & Carroll, 2001). Furthermore, maximum likelihood estimation with MLR provides standard errors and Chi^2 test statistics robust to non-normality and non-independence of observations when used with TYPE = COMPLEX (L. K. Muthén & Muthén, 1998-2017, p. 668).

[48] On the differences between the performance of WLSMV and MLR in Mplus see the relevant literature (e.g., Beauducel & Herzberg, 2006; Li, 2016; Scherer, 2020).

Comparing the functionality of WLSMV and MLR in confirmatory factor analysis, Li (2016) showed that MLR outperforms WLSMV when data is not normally distributed (Li, 2016). Lastly, MLR estimates a linear regression for continuous data and a logistic regression for categorical data, and thus, appropriately accounts for categorical dependent variables when numerical integration is applied.

Particular to this study is the use of categorical data to fit measurement models of parental involvement that function as dependent and independent constructs in the SEM. On that note, Svetina et al. (2020) argued that the acknowledgment of the data being categorical is crucial in terms of obtaining correct parameter estimates, and thus, ensuring reliable model fit and cross-group comparisons when estimating measurement models through confirmatory factor analysis (Svetina et al., 2020). Nevertheless, categorical dependent variables utilized to fit the measurement model of this study are treated as continuous instead, and WLSMV estimation as well as numerical integration are disregarded. This is because acknowledging the complex structure of the data and the large sample size, numerical integration becomes computationally demanding (L. K. Muthén & Muthén, 1998-2017, p. 758). Furthermore, there is a strong argument for making a distinction between the number of answering categories in categorical items, and to assume that not all categorical data needs to be treated equally. For instance, Robitzsch (2020) argued that it is justifiable to fit a linear model for categorical items with at least three answering categories, as long as robust maximum likelihood estimation (MLR) is applied. Problematic, however, are dichotomous items, which should not be treated as continuous information (Robitzsch, 2020). Dichotomous items are not included in the model of this study, and hence, the application of MLR neglecting numerical integration is considered appropriate to accommodate for the nature of the data.

Fit indices

> "*Fit indices yield information bearing only on the model's lack of fit. More importantly, they can in no way reflect the extent to which the model is plausible; this judgment rests squarely on the shoulders of the researcher.*" (Byrne, 2012, p. 77)

Although SEMs are heavily reliant on a strong and substantive theoretical framework (e.g., Byrne, 2012; Marsh et al., 2004; Nagengast & Marsh, 2014), researchers commonly draw on different goodness-of-fit statistics to reinforce judgment on whether a model fits the data or not. The following section touches briefly on the subject, focusing on such indices that are relevant to this study. However, a more in-depth take on the topic is provided by Kline (2016), Byrne (2012), Marsh et al. (2004), and Hu & Bentler (1999) among others.

Chi^2

First and foremost, the Chi-Square test statistic, which — in this context — refers to the Chi-Square (also represented as Chi^2 or X^2) statistic of the *Likelihood Ratio Test*, has been the most broadly used concept for the evaluation of model fit in SEM. The Chi^2 test statistic represents to what extent the null hypothesis (H_0), which assumes the parameters of the postulated model to be valid, is true. More precisely, Byrne (2012) states the following:

> "The probability value associated with X^2 represents the likelihood of obtaining a value that exceeds the X^2 value when H_0 is true. Thus, the higher the probability associated with X^2, the closer the fit between the hypothesized model (under H_0) and the perfect fit." (Byrne, 2012, p. 67)

Hence, a significant Chi^2 test implies that the null hypothesis H_0 needs to be rejected (Geiser, 2011). Unfortunately, Chi^2 has shown to be very sensitive to sample size (e.g., G. W. Cheung & Rensvold, 2002; Svetina et al., 2020), and thus, models often need to be rejected (Byrne, 2012).

Another particularity associated with the computation of the Chi^2 results from the application of the MLR estimator. Given that MLR accounts for non-normally distributed data, the commonly known Chi^2 Difference Test is actually not Chi^2 distributed. To account for that, Satorra (2000) and Satorra & Bentler (2010) provided a scaling correction factor to achieve an improvement of the Chi^2 approximation in order to be used for the

Satorra-Bentler scaled Chi^2 Difference Test (Satorra & Bentler, 2010, p. 243). The application requires two equations (presented after Kleinke et al., 2017; Satorra & Bentler, 2010).[49]

First, the *difference test scaling correction* (cd) is calculated based on the scaling correction factor (c0), and the degrees of freedom (d0) of the more restrictive model (i.e., the nested model); and the scaling correction factor (c1) as well as the degrees of freedom (d1) of the less restrictive model (i.e., the comparison model):

$$\text{cd} = (\text{d0 x c0 - d1 x c1}) \; / \; (\text{d0 - d1})$$

Second, the Satorra-Bentler scaled Chi^2 Difference Test is applied, resulting in the difference TRd, whereas T0 represents the Chi^2 value of the nested model; and T1 represents the Chi^2 value of the comparison model, both obtained from the Mplus output:

$$\text{TRd} = (\text{T0 x c0 - T1 x c1}) \; / \; \text{cd}$$

As a result, the corrected difference (TRd) is nearly normally distributed. The significance of the corrected difference needs to be estimated separately.[50]

Given that this study applies the MLR estimator, a correction of Chi^2 according to the Satorra-Bentler scaled Chi^2 Difference Test is in order. However, as pointed out by Feld (2018), the option is unavailable for imputed data, and thus, for data from the prePIRLS assessment, given that analyses thereof require the imputation of plausible values (Feld, 2018).

Although, this is true for the SEM, confirmatory factor analyses and multi-group measurement invariance testing are based on only one data set in order to obtain modification indices and identify the appropriate level of measurement invariance (sec. 8.2.2) . Hence,

[49] Difference testing with the Satorra-Bentler scaled Chi^2 Difference Test can be performed by using either Chi^2, or the likelihood. Both are provided in the Mplus output and can be used interchangeably leading to the same results (Asparouhov, 2017). Details on the corresponding steps and equations are provided on the Mplus website at: https://statmodel.com/chidiff.shtml

[50] For this study, the following application has been used based on the recommendations of Kleinke et al. (2017) at: https://www.fourmilab.ch/rpkp/experiments/analysis/chiCalc.html

respective analyses are not subjected to multiple imputation, and therefore allow for the performance of the Satorra-Bentler scaled Chi^2 Difference Test. As for the full SEM, all five data sets are imputed, and hence, the corrected test is not applicable. Instead, model fit indices are used to evaluate the model fit.

TLI/CFI, RMSEA and SRMR

Different fit indices have been developed, aiming to overcome pertinent issues associated with the sensitivity of Chi^2. There are two groups of fit indices, that are incremental (i.e., comparative) and absolute indices. Though, the former tend to be more widely applied in SEM (Byrne, 2012). The difference between the two is the following:

> "Whereas incremental indices of fit measure the proportionate improvement in fit of a hypothesized model compared with a more restricted, albeit nested, baseline model (Hu & Bentler, 1999), absolute indices of fit assess the extent to which an a priori model reproduces the sample data." (Byrne, 2012, pp. 70–71)

The most common incremental fit statistics are the *Comparative-Fit-Index* (CFI) and the *Tucker-Lewis-Index* (TLI). The CFI is based on the comparison of a hypothesized model with a less restrictive baseline model, which is freed from assuming covariances between observed variables (Byrne, 2012; Geiser, 2011). The CFI is set to a scale of 0 to 1, with a value closer to 1 indicating a well-fitting model. Although functioning on the same logic as the CFI, the TLI can take values above and beyond the range of 0 to 1. Therefore, the TLI is occasionally also classified as a non-normed index (Byrne, 2012). Both CFI and TLI should meet a threshold of .95 or .97 in order for a model to be assumed well-fitting to the data (Geiser, 2011). However, acceptable fit is achieved when CFI and TLI meet a threshold of at least .90 (Nagengast & Marsh, 2014).

Popular absolute fit statistics are the *Root-Mean-Square-Error-of-Approximation* (RMSEA) and the *Standardized-Root-Mean-Square-Residual* (SRMR). The RMSEA is a measure of approximate data fit expressed by the degrees of freedom, and thus, sensitive to model complexity (Byrne, 2012). The index has been claimed as a routine measure of

model fit, given that it is reliably sensitive to model misspecifications (Hu & Bentler, 1999). Furthermore, the SRMR is the standardized value of the *Root Mean Square Residual*, which "represents the average residual value derived from the fitting of the variance-covariance matrix for the hypothesized model (...) to the variance-covariance matrix of the sample data" (Byrne, 2012, p. 76). As opposed to incremental fit indices such as CFI and TLI, absolute fit indices decrease as model fit improves. Therefore, a RMSEA and SRMR at or below .05 is assumed to represent a well-fitting model (Geiser, 2011). However, some sources also argue for an RMSEA at or below .06 to indicate good fit, and values at or below .08 to indicate acceptable fit (Hu & Bentler, 1999; Nagengast & Marsh, 2014).

The choice of suitable goodness-of-fit statistics is often troublesome, given the variety of options and their individual particularities with regard to sample size, model complexity, and assumptions of data distribution (Byrne, 2012). Therefore, it is strongly recommended to acknowledge several fit indices in order to evaluate model fit (Byrne, 2012; X. Fan & Sivo, 2005; Hu & Bentler, 1999; Kline, 2016). On that note, Hu & Bentler (1999) published a widely acknowledged article on the issue of cut-off criteria for fit indices in SEM (Hu & Bentler, 1999), and proposed a hands-on solution to the issue. Reviewing several goodness-of-fit indices, they came to the conclusion that the application of a combinational rule or *two-index strategy*, that is the combination of differently functioning but complementary indices, such as CFI and SRMR, is recommended to assess model fit reliably (X. Fan & Sivo, 2005; Hu & Bentler, 1999). Furthermore, Hu & Bentler (1999) proposed cut-off criteria that should be recognized when the combinational rule is applied (Hu & Bentler, 1999). Although respective recommendations of thresholds have been critically reviewed and mostly invalidated (e.g., X. Fan & Sivo, 2005), a combination of goodness-of-fit indices is still considered an appropriate procedure. For example, Kline (2015) suggested the following combination of indices to be considered a standard for model fit evaluation. Recommended fit indices are Chi^2, RMSEA,[51] CFI, and SRMR. By

[51] The RMSEA is commonly accompanied by 90% confidence intervals (Kline, 2016). In this study, only the RMSEA is printed in the results. Confidence intervals may be obtained form the application of

doing so, both groups of incremental and absolute indices are considered. Hence, model evaluation is based on two perspectives, namely the comparison of a hypothesized model to a less restrictive baseline model, and the application of a hypothesized model with regard to the question of how well it represents the sample data.

Although, aimed at approaching model fit evaluation from a sophisticated and comprehensive perspective, Kline (2012) emphasized the limits of fit indices, and hence, related to the opening statement of this section. Pointing towards the requirements for investigating causal inferences between constructs, he determines that in behavioral or social sciences, respective requirements are almost always violated by design. This is because in those fields of research and with the data at hand often being cross-sectional, one rarely knows the true causal model, but still hypothesizes and tests thereof (Kline, 2012). On that note, Kline (2012) derived that fit indices can only support the assumption whether a model is consistent with the data. "In this way, SEM can be seen a disconfirmatory technique, one that can help us to reject false models (...), but it basically never confirms a particular model when the true model is unknown" (Kline, 2012, p. 114). Respective conclusions on the explanatory power and validity of SEM as well as fit indices need to be considered for the evaluation of results.

Summary

To summarize, CFI and TLI ($>.97/.95/.90$) as well as RMSEA ($<.05/.06/.08$) and SRMR ($<.05$) are reported and considered in this study. As elaborated, Chi^2 is not heeded for the evaluation of model fit of the full multi-group SEM, though consulted for the evaluation of measurement invariance (sec. 8.2.2). Lastly, fit indices are acknowledged for the sake of model evaluation. Thus, they are not treated as the ultimate tool for neglecting or accepting a model.
They are rather seen as an additional tool alongside other relevant aspects, such as a

the Mplus syntax included in the extended appendix (sec. 'Measurement Invariance Testing (Mplus Syntax)' and sec. 'Structural Equation Model (Mplus Syntax)').

substantial theoretical framework, state of research, and properties of the data itself, to determine the validity of the model.

8.2.2. Multi-Group Invariance Testing

> "*Given the complex relations between the sensitivity of measurement tests invariance, degrees of noninvariance, and potential bias in group comparisons, it is difficult to propose statistical standards for testing measurement invariance.*" (F. F. Chen, 2007, p. 501)

A pivotal aspect of this study is the comparison of parental involvement and its effect on student performance among different groups of the South African population, which are socioeconomic low- and high-performers (sec. 8.2.4).

When thinking about invariance, one needs to distinguish two overarching constructs, each focusing on a different aspect of invariance inherent to the method of SEM. In that regard, Little (1997) introduced category one and category two invariance, and Dimitrov (2010) further elaborated on category one invariance as measurement invariance, and on category two invariance as structural invariance. As indicated by the terminology, measurement invariance tests for invariance of the measurement model between groups. Structural invariance, on the other hand, focuses the invariance of factor variances and covariances between groups (Dimitrov, 2010). Hence, in SEM, measurement invariance is a prerequisite for structural invariance (G. W. Cheung & Rensvold, 2002).

The following section primarily emphasizes procedures subjected to measurement invariance. However, since structural invariance follows the same premises, especially with regard to model fit, the topic is revisited in section 8.3.

Concepts of Measurement Invariance

In order to ensure comparability of the latent factorial structure of parental involvement between groups, the invariance of the measurement needs to be ensured. Measurement invariance represents the comparability of the measurement model, namely the latent factor structure, between groups. Testing model parameters such as factor loadings, intercepts of manifest variables, and the residual variance, the aim of measurement invariance is to

prove that respective measurement models behave invariant (i.e., similar) between groups (Kleinke et al., 2017; Millsap & Olivera-Aguilar, 2012). Thus, violations of measurement invariance would entail that differences or similarities between groups are merely random or caused by other influential factors, but not due to the latent factor structure that is observed. "In that case, group comparisons on the measured variables would have uncertain interpretations" (Millsap & Olivera-Aguilar, 2012, p. 381).

Based on pioneering work by Meredith (1993), three levels of measurement invariance are acknowledged. The first and least restrictive level is configural invariance. Configural measurement invariance looks at the equality of factorial structures and is established when the number and patterns of factor loadings are equal between groups. The second level is referred to as metric measurement invariance, also known as weak factorial or weak measurement invariance. It is established when the requirements for configural invariance are fulfilled, and factor loadings are found to be equal between groups. The third and most restrictive level of measurement invariance is known as scalar or strong factorial invariance. It is given if requirements of level one and two are fulfilled, and intercepts of manifest variables are proven to be identical.[52] Levels of measurement invariance are hierarchically ordered, meaning that configural invariance needs to be established in order for metric invariance to be considered, and metric invariance needs to be established in order for scalar invariance to be considered (Kleinke et al., 2017).

The establishment of measurement invariance can be realized through a *Step-Down* or a *Step-Up* approach. For both options, multiple-group comparison is applied in order to compare measurement models with deviating levels of restrictiveness. The Step-Down approach starts with the most restrictive model and gradually loosens restrictions of equality to identify the appropriate level of measurement invariance. Vice versa, the Step-Up approach starts with the least restrictive model, assuming configural invariance, and introduces increasingly restrictive models to be compared to the original, least restrictive

[52] Additionally, the literature recognizes strict factorial invariance, given when also residuals of the manifest variables are found to be equal across groups. However, this level is rarely considered in applied research (Kleinke et al., 2017), and hence, untended in this study.

model. If the change in model fit parameters does not exceed agreed upon thresholds, the respective level of measurement invariance can justifiably be assumed (Kleinke et al., 2017). To conclude, the Step-Up approach is known to be the more commonly applied procedure (Kleinke et al., 2017). Thus, it is used in this study as well.

Based on Brown (2015), Kleinke et al. (2017) suggested the following steps towards establishing measurement invariance within the scope of a Step-Up approach. To start, the fit of the measurement model should be checked separately for each group. Kleinke et al. (2017) refer to this step as the establishment of configural invariance. Next, the baseline model should be built, that is, a multiple-group comparison of the measurement model. Being least restrictive, the baseline model is a model of independence, assuming that there is no covariance between the variables of interest (Geiser, 2011). In the baseline model, all parameters of the measurement model are estimated freely, meaning that they are allowed to differ in terms of factor loadings, intercepts, and latent means (Kleinke et al., 2017). Thus, the baseline model serves as a reference model, as to which the more restrictive models of metric and scalar invariance are compared to.
Each level of measurement invariance enables different options for multiple group comparison. With configural invariance established, it is demonstrated that the factorial structure as well as patterns of factor loadings are equal (Geiser, 2011). Metric invariance needs to be established in order to justifiably compare factor loadings between groups, and thus, make assumptions about similarities and differences in structural paths. Furthermore, establishing factorial or scalar invariance is required to compare latent means, based on the invariance of intercepts of manifest items. Lastly, strict factorial invariance is reached when, in addition to all aforementioned conditions, also measurement errors or residuals hold to be invariant between groups (Kleinke et al., 2017). However, as aforementioned, this study does not account for strict factorial invariance.

Partial Measurement Invariance

Although administered in a standardized way within and across education systems participating in prePIRLS 2011, items of parental involvement, and latent factors resulting therefrom, might not measure the same thing across different groups. Given that the South African population is highly segregated in terms of socioeconomic status and language patterns among others, there is reason to expect differences in the way that items on parental involvement have been received and answered within different groups of the population. To account for that, partial measurement invariance needs to be considered. Partial measurement invariance comes into play when metric and scalar invariance are not reached, given that restrictions of factor loadings and intercepts have been introduced to the measurement model. Addressing the issue of non-invariance, Byrne et al. (1989) raised the option of partial measurement invariance (Byrne, 2012). In short, the procedure allows for the gradual release of factor loadings, which have been fixed but did not reach a certain level of invariance. When partial measurement invariance is successfully established, factor loadings of concern are left non-invariant, whereas other factor loadings remain invariant. Thus, comparability of measurement and path models between groups may still be ensured, even though certain items of a construct are considered non-invariant between groups (Byrne et al., 1989; Kleinke et al., 2017).

However, Byrne et al. (1989) strongly claimed "the procedure to be most effective with minimal model modifications; the relaxation of many parameters is likely to yield an unsuccessful cross-validation" (Byrne et al., 1989, p. 465). Revisiting the issue and referring to Byrne et al. (1989), Dimitrov (2010) presented analyses on measurement and structural invariance, and found that relaxing up to 20% of the fixed parameters is an acceptable threshold in applied research (Dimitrov, 2010). On that note, Cheung & Rensvold (2002) concluded that non-invariant items will not meaningfully compromise group comparisons when only few items are freed from restrictions (G. W. Cheung & Rensvold, 2002).

Evaluating Measurement Invariance

To evaluate whether the respective levels of measurement invariance hold for the tested measurement models, differences in fit statistics are commonly considered.

Traditionally, the Chi^2 Difference Test is conducted to evaluate whether models of the same parameters are invariant, that is equal to each other across groups. However, the Chi^2 Difference Test has been shown to be rather sensitive. Thus, additional fit statistics are recognized as important complements to Chi^2. Relevant work in this regard has been provided by Cheung & Rensvold (2002), Chen (2007), and Rutkowski & Svetina (2014) among others.[53] Most commonly, differences in the CFI and RMSEA between the baseline model and more restrictive models are consulted (Svetina et al., 2020).

As for CFI, Cheung & Rensvold (2002) suggest tolerating a maximum decrease of -.010 in order to establish a stricter level of measurement invariance (G. W. Cheung & Rensvold, 2002). Recall, a decrease in CFI implies worse model fit. Along those lines, Chen (2007) posited that when testing metric invariance compared to the baseline model, a CFI decrease no greater than -.010 should be accompanied by an increase in RMSEA no greater than .015 or SRMR no greater than .030 in order for metric invariance to be justifiably assumed in groups of n =/> 300. Recall, an increase in either RMSEA or SRMR implies worse model fit. Moving to the more strict model of scalar invariance, increases in SRMR should not exceed .010, while guidelines for changes in RMSEA and CFI remain the same as when testing for metric invariance against the baseline model. Based on two Monte-Carlo studies, Chen (2007) furthermore summarized that CFI is the index of choice, given that RMSEA and SRMR are at times affected by sample size and model complexity (F. F. Chen, 2007).

Addressing this issue among others, Rutkowski & Svetina (2014, 2017, 2020) presented a series of studies investigating fit indices for measurement invariance within different

[53] The herein listed publications focus the investigation of measurement invariance in models assuming normally distributed data. More recent research, for example by Svetina et al. (2020), focuses the investigation of measurement invariance of models built on categorical data. Although items used in this study are categorical by nature, they are treated as non-normally distributed continuous items instead.

settings. Taking into account large sample sizes when comparing more than two groups, Svetina et al. (2020) argued for less strict thresholds in fit indices when evaluating levels of measurement invariance based on data with the aforementioned characteristics (Svetina et al., 2020). On that note, they call for a decrease in CFI up to -.020 in order for metric invariance to hold, and for a decrease up to -.010 in order for scalar invariance to hold. Increases in RMSEA should not exceed .030 for metric invariance and .010 for scalar invariance (Rutkowski et al., 2014). Looking at leading studies in the field of internationally comparative assessments in education, Rutkowski & Svetina (2014) find that a formal investigation of measurement invariance is provided for some studies, for example the *Teaching and Learning International Survey* (TALIS) and the *Programme for International Student Assessment* (PISA). In those studies, the basis for evaluating measurement invariance are CFI and RMSEA (Rutkowski et al., 2014).

Summary and Implications for the Study

To summarize, multi-group invariance testing using a Step-Up approach is considered in this study along the lines of quality assurance and accuracy of empirical practice. To that end, measurement invariance of the latent construct is investigated first, followed by checking for invariance of the structural paths. However, given that the groups of this study represent highly segregated entities of the South African population deviating in ethnic, cultural, and socioeconomic backgrounds, finding non-invariance of measurement for the construct of parental involvement is considered likely. That is, this study assumes that parental involvement and the importance thereof might have been perceived differently between groups. In this case, partial measurement invariance is considered.[54]

[54] Conveniently, Mplus offers an automated approach to test for measurement invariance on configural, metric and scalar level. Respective models are compared using the Chi^2 Difference Test. Although advantageous in terms of quickly providing correct model results for each level of measurement invariance, the automated approach does not provide separate fit indices other than Chi^2 and corresponding p-values. Furthermore, the automated approach does not allow testing for partial measurement invariance, and is incompatible with multiple imputation. Therefore, the Mplus automated option to test for measurement invariance is merely used for the exploration of the data, and for undocumented preliminary analysis.

Lastly, this study primarily considers changes in CFI and RMSEA to evaluate levels of measurement and structural invariance. Given that this study compares two groups, and deals with sample sizes of n > 2000, thresholds of fit statistics are applied according to Chen (2007) and Rutkowski & Svetina (2014), if indicated. Furthermore, Chi^2 as well as results of the Chi^2 Difference Test are reported for further reference. Though, given that the data is estimated using the MLR estimator, the scaling correction factor of Chi^2 is used, and the Satorra-Bentler scaled Chi^2 Difference Test is applied. However, detecting false model misfit through Chi^2 is expected, given that the index is known to be sensitive to sample size and leading to "inflated type I error rates" (Svetina et al., 2020, p. 114). Hence, all three indices are considered in conjunction, keeping in mind the benefits of the combinational rule after Hu & Bentler (1999). Generally, recommendations on the interpretation of model fit indices as well as corresponding thresholds should not be treated as "golden rules" but with caution (Marsh et al., 2004; Nagengast & Marsh, 2014).

8.2.3. Selection of Variables

In the following section, the items used to operationalize parental motivation, home-based parental involvement, and student outcomes in learning are presented. All items stem from the home- and student-background questionnaire of prePIRLS 2011. It must be noted that items utilized for this study are scarcely continuous, but mostly categorical consisting of four answering categories. Therefore, all metric scales of latent constructs in PIRLS and prePIRLS are artificial in nature since they do not result from true numeric items. This is particularly important for the interpretation of averages, and the derivation of groups.

Parental Motivation for Involvement

Parents' motivation for educational involvement is defined as the pre-decisional phase of the model of human action behavior. That is, it refers to the motives and motivating incentives that lead parents to get involved in their children's education.

As described in section 3.3, Walker et al. (2005) suggested several motives for parental involvement. In this study, the item of parents feeling included in their children's education by the school (Do you think your child's school includes you in your child's education? (*ASBH10A*)) (Foy & Drucker, 2013a) is used as a proxy for the dimension of *Parents' Perceptions of General School Invitations* (Walker et al., 2005).[55] The original scale on *Perceptions of General School Invitations* by Walker et al. (2005) consists of six items, namely parental perceptions of teachers being interested in cooperation, schools being welcoming, availability of parent activities at the school, information on meetings and special school events, schools making an effort to contact parents immediately in case of an issue with the child's education, and teachers being informative regarding the child's educational progress (Walker et al., 2005, p. 101). The operationalization of the original scale was based on a measure proposed by Griffith (1996) regarding parents' satisfaction with schools, and yielded a scale reliability of .88 (Walker et al., 2005). Hence, the item on parents' perception of being included to their children's education by the school is considered to be in line with the content of the original scale. In addition to *Parents' Perception of Invitations for Involvement from Others*, Walker et al. (2005) appointed several other motivational aspects of parental involvement (e.g., *Parents' Motivational Beliefs* and *Parents' Perceived Life Context*, as well as dimensions thereof). However, given the limited availability of items in prePIRLS 2011, this study can only account for the aforementioned dimension of *Parents' Perceptions of General School Invitations*.

As argued in section 4.2.1, parental educational aspirations are considered to be a motivational factor for parental involvement as well. Hence, this study applies parental educational aspirations (How far in his/her education do you expect your child to go? (*ASBH18*)) (Foy & Drucker, 2013a). This is based on theoretically guided assumptions,

[55] Items on self-perceived school efforts to include the parents (Do you think your child's school should make a greater effort to include you in your child's education? (*ASBH10B*)), and self-perceived school efforts to keeping parents informed (Do you think your child's school should do better at keeping you informed of his/her progress? (*ASBH10E*)) (Foy & Drucker, 2013a) have been dismissed as possible dimensions for parental motivation in this study, given their weak statistical performance for the South African population.

and empirically based findings emphasizing the importance of the concept for motivating parental involvement (e.g., Castro et al., 2015; X. Fan & Chen, 2001; Jeynes, 2007; Peng & Wright, 1994; Seyfried & Chung, 2002; K. Singh et al., 1995; Wilder, 2014; Xu et al., 2010).

Home-Based Parental Involvement Activities

Along the lines of this study's theoretical framework, forms of parental involvement, and more specifically home-based parental involvement, represent the volitional phase of the model of action behavior.

Referring back to empirical work by Keith et al. (1993), and Singh et al. (1995) among others (sec. 3.3), Fan & Chen (2001) argued that "parental involvement subsumes a wide variety of parental behavioral patterns and parenting practices" (X. Fan & Chen, 2001, p. 3), and thus, "different dimensions of parental involvement should be measured separately instead of being summed up into a general composite" (X. Fan & Chen, 2001, p. 17). This is to allow for a more defined effect of parental involvement on student outcomes in learning, compared to a unidimensional construct (X. Fan & Chen, 2001).
On that note, Walker et al. (2005) argued for a multidimensional concept of parental involvement as well, though accounting for the common distinction between home-based and school-based involvement forms. With regard to home-based involvement, they proposed a five-item scale, including questions on whether someone in the family is doing the following on a regular basis: talk with the child about the school day, supervise the child's homework, help the child study for tests, practice a subject with the child, and read with the child (Walker et al., 2005). School-based involvement, on the other hand, is based on questions regarding the willingness to help out at the child's school, attend special events at the school, volunteer to go on class field trips, attend PTA meetings, and go to the school's open houses (Walker et al., 2005). Both concepts are based on previous work by Epstein & Salinas (1992), arguing that home-based and school-based parental involvement activities are an essential part of children's successful educational upbringing

(Epstein & Salinas, 1992). In summary, Walker et al. (2005) found that motivational factors on level one explained 33% of the variance in home-based involvement activities, and 19% of the variance in school-based involvement activities (Walker et al., 2005).

In prePIRLS, information on both home-based and school-based involvement activities are provided.
Home-based involvement is reflected by a battery of eight items similar in content to the aforementioned selection of items presented and empirically tested by Walker et al. (2005). Respective items are included to the home-background questionnaire, and therefore answered by the parents themselves. Hence, such as the proxy for parents' perceptions of invitations for involvement from the school, home-based parental involvement forms are self-reported. On the contrary, information on school-based parental involvement are collected from principals and teachers. Given that this study focuses parents' motives and actions, only self-perceived home-based involvement activities are included to the SEM. Using data from prePIRLS, and including school-based involvement would entail both, an extension of the theoretical framework by an additional perspective, namely school-based parental involvement, principals' and teachers' perception thereof, and the application of a multi-group, multilevel SEM. Hence, the inclusion of respective measures would exceed the scope of this study. Hence, they are left for future analyses. However, in line with the theoretical framework, this study operationalizes home-based parental involvement through the frequency of the following parental actions (Foy & Drucker, 2013a):

- Discussing schoolwork with the child (*ASBH09A*)
- Helping the child with homework (*ASBH09B*)
- Ensuring that the child sets aside time for homework (*ASBH09C*)
- Asking the child what it has learned in school (*ASBH09D*)
- Checking the child's homework (*ASBH09E*)
- Help the child to practice reading (*ASBH09F*)
- Help the child to practice math (*ASBH09G*)

- Talk with the child about what the child reads (*ASBH09H*)

Results of exploratory factor analysis on respective items lead to a two-factor solution on home-based parental involvement. Factors are parental *Supervision* and parental activities of *Practice*. Hence, rather than opting for the common distinction of home-based and school-based parental involvement forms, this study applies a subdivision of home-based parental involvement, and thus, accounts for the need of more defined measures of the construct (X. Fan & Chen, 2001).

Student Outcomes

Lastly, the evaluation of student outcomes represents the post-actional phase of the model of action behavior.

Against the background of the theoretical framework presented in section 3.3, student outcomes in learning are first and foremost defined as student performance in certain subjects, for example math and reading. The prePIRLS assessment offers several scales on reading literacy. In this study, all five plausible values of the scale on overall reading achievement (*ASRREA01 - 05*) are utilized.

Looking at Fan & Williams (2010), student outcomes are furthermore understood as student engagement, student self-efficacy, and student intrinsic motivation (sec. 4.2). Whereas student engagement is not sufficiently represented in the data, student self-efficacy and intrinsic motivation are operationalized by scales on students' confidence in reading (*ASBGSCR*) (Martin & Mullis, 2012b), and students' self-report on whether they like to read (*ASBGSLR*) (Martin & Mullis, 2012c). Both scales are included in the model of this study as additional measures of student educational outcomes. Originally based on students' responses to several statements regarding questions on self-perceived cognition of the ability to read, and attitudes towards reading, the scales were created using IRT scaling methods, and the Rasch partial credit model after Masters & Wright (1997) (Martin et al., 2012). Given the elaborate scaling procedure, and to ensure comparability of the results to other studies working with the same constructs, the scales are included as continuous constructs to the model of this study. Though, this is at the

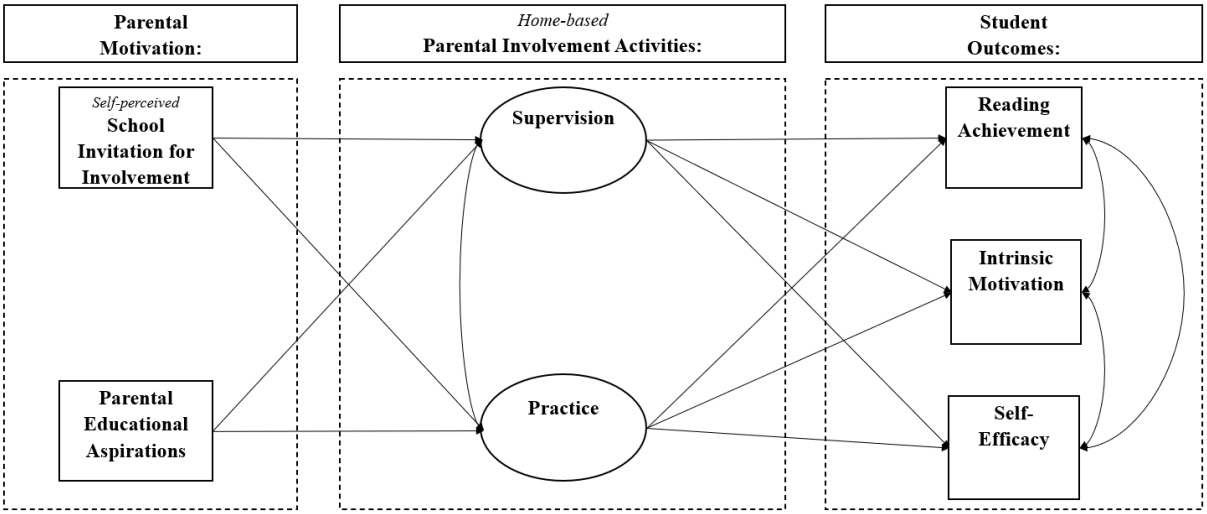

Figure 13: Blueprint of the SEM based on the theoretical framework of this study, including selection of variables. Own illustration.

costs of measurement accuracy as well as a matching available scales with the original constructs of students' self-efficacy as well as students' intrinsic motivation as applied by Fan & Williams (2010).

By including overall student performance in reading as well as additional educational outcome measures, namely students' self-efficacy and students' intrinsic motivation, the effect of home-based parental involvement activities is evaluated regarding several outcomes rather than one. This is in line with the state of research, suggesting that parental involvement is particularly important to students' personal motivational beliefs in addition to overall performance (Eccles et al., 1998; W. Fan & Williams, 2010; Grolnick & Slowiaczek, 1994).

Summary and Implications for the Study

In this study, SEMs are build along the lines of the theoretical framework (sec. 3), and empirical state of research (sec. 4), utilizing the set of variables discussed in this section. Transferring the premises of SEM (sec. 8.2.1), figure 13 mirrors the theoretical framework and shows the blueprint for the empirical model of this study, including the variables and constructs presented in this section.

Along the lines of SEM notation (sec. 8.2.1), linear paths map a directional association between the motivation for parental involvement with parental involvement activities. In accordance with the state of research as well as derived hypotheses, respective paths account for assumptions posited for the association between parental motivation, parental involvement, and the evaluation thereof. On that note, the blueprint for the SEM also accounts for the assumed correlations between dimensions of parental involvement, as well as the dimensions of student outcomes. Lastly, latent and observed constructs are indicated by circles and squares.

8.2.4. Measures of Group Identification

A central aspect of this study is the distinction between societal groups, accounting for the segregation and unequal living opportunities caused by Apartheid policies of the past. Hence, the SEMs of this study are estimated for several (pairs of) groups of the South African population. Thus, rather than including proxies of socioeconomic status as contextual information to the SEM itself, the distinction of socioeconomically different groups allows for a separate exploration of the model by groups and pairs of groups.

Of particular interest are socioeconomic low-performers, and thus, they are considered the reference group of this study. However, additional groups such as socioeconomic high-performers are considered as well, aimed at providing a counterpart to the reference group. Based on the reflections of the population's constitution and characteristics (sec. 2), several aspects are considered relevant for the identification of socioeconomic status in South Africa. Hence, to identify respective groups, different variables are utilized.

First, incorporating information in line with economic, cultural and social capital (e.g., B. B. Bernstein, 1975; Bourdieu & Passeron, 1977; Coleman, 1988), the *Home Resources for Learning* (HRL) scale functions as a well-established proxy for one's socioeconomic status (e.g., Caro et al., 2014). Second, information such as area of living and language spoken at home are considered to be decisive characteristics for the distinction between socioeconomic low- and high- performers in South Africa as well. Although data of prePIRLS 2011 does not provide information on the area of living, information on the

	n_{Valid}	$\%_{Valid}$	$n_{Missing}$	$\%_{Missing}$
Groups of Home Resources for Learning				
LHRL (25th Percentile)	2863	18.2		
HHRL (75th Percentile)	2524	16.0		
Others	3831	24.2		
MHRL (Missing)			*6544*	*41.6*
Language Groups				
Native	9093	57.8		
NonNative (English or Afrikaans)	3040	19.3		
Others	1058	6.7		
Missing Values			*2553*	*16.2*

Note:
Group identification variables are derived from multiple items. Home Resources for Learning are derived from the scale *ASBGHRL;* and Language Groups are identified through the computation of a new variable based on *ITLANG* and *ASBG03*.

Table 1: Descriptive properties of groups. Own illustration, based on own analysis using the South African data set of prePIRLS 2011.

language environment that students are primarily exposed to at home and at school are available.

Aimed at accounting for respective dimensions of socioeconomic status, groups are based on the HRL scale, as well as information on language spoken at home and at school. Table 1 presents the descriptive properties of all groups, summarizing unweighted valid and missing information in absolute (n) and relative numbers (%).

Groups by Home Resources for Learning

The HRL scale consists of five categorical items providing information on number of books in the home reported by the students, number of children's books in the home reported by the parents, number of home study supports reported by the students, as well as highest level of education and highest level of occupation of either parent reported by the parents themselves. The prePIRLS data set offers a categorical index variable as well as a continuous scale on HRL. For the index variable, cases with a score of 7.3 or below on the HRL scale were summarized in the lowest category *few resources for learning* and cases with a score of 11.9 and above were summarized in the highest category *many resources for learning*. To allow for comparability with other studies utilizing prePIRLS data, the HRL scale is used for this study. Though, rather than utilizing the scale as such, the 25^{th} and the 75^{th} percentile of the distribution are used to divide the HRL scale into 25% of South Africans who showed the lowest and highest socioeconomic resources according to

information summarized to the scale. Respective percentiles are chosen to ensure groups that are large enough to allow for a well-functioning SEM.

Descriptive statistics on the HRL scale show large shares of missing information, which applies to over 40% of the cases. Although it is not uncommon to detect large shares of missing data in variables regarding socioeconomic status, almost half of the cases missing information thereon is considered excessive, especially with regard to the threshold of 30% missingness to be tolerable in empirical research (Wirtz, 2004). Thus, rather than estimating missing information in HRL through procedures of missing data estimation (sec. 8.2.5), this study opts for dealing with those who have missing information on the HRL scale by summarizing them to an additional group. In doing so, the multi-group SEM of this study accounts for three groups of HRL, which are those who have access to few resources (group LHHRL), those equipped with many resources (group HHRL), and those whose information on HRL are missing (group MHRL). Furthermore, it needs to be noted that although the 25^{th} and the 75^{th} percentile are used as thresholds for group identification, group sizes do not equal precisely 25% and 75% of the distribution. That is because the items, which the scale has been derived from, are categorical in nature, and can therefore not result in exact numbers.

Language Groups

In addition to home resources for learning, language is an indicator associated with socioeconomic status and life opportunities in South Africa as well (sec. 2.2.1).
The country's Black-African population most frequently speaks a native African language (i.e., isiZulu and isiXhosa) while also showing lower socioeconomic status (e.g., lower educational and occupational status). Furthermore, it was shown that those taught in a native African language were also exposed to lack of school resources for learning (e.g., Howie et al., 2012) and that the majority of students with access to many home resources for learning were taught in English or Afrikaans. On the contrary, only a few students taught in a native African language had access to a comparable set of resources (van

Staden et al., 2016). Therefore, it is not surprising that student outcomes in learning were found to diverge noticeably between language groups. For example, students educated in English or Afrikaans outperformed students educated in native African languages (van Staden & Howie, 2014; van Staden et al., 2016).

Relevant to respective findings is a particularity of the South African education system, namely the discrepancy between students' home language and the language of instruction. As shown in section 2.2.4, the South African *Language in Education Policy* stipulated that students are educated primarily in non-native languages, namely English or Afrikaans, from grade 4 onward. The policy inevitably resulted in students being educated and tested in a language that might not be their home language. Aimed at assessing the effects of respective discrepancies, a study by van Staden et al. (2016) based on prePIRLS 2011 showed that student performance in reading decreased when the language of instruction (i.e., the language of test) was different from the home language (van Staden et al., 2016).

Aiming to achieve homogeneous groups for the analysis of parental educational investment behavior among South African socioeconomic low- and high-performers, the presented state of research indicated two things. First, the twofold pattern of language groups needs to be acknowledged. Second, the discrepancy between home language and language of instruction needs to be accounted for. On that note, this study utilizes a variable indicating the language of test (*ITLANG*) and compares it to the frequency of language of test spoken at home (*ASBG03*), resulting in two homogeneous language groups. These are labeled *Native* (i.e., native African languages) and *NonNative* (i.e., English or Afrikaans). Language of test (*ITLANG*) is a tracking variable, derived from tracking forms that collect information from students, teachers, and schools (Foy & Drucker, 2013b). Given that the South African prePIRLS assessment was administered to fourth-grade students in one of 11 official languages acknowledged by the South African government, the variable consists of 11 categories. There are no missing cases in the tracking variable, providing the most complete set of information with regard to the language environment that students were

primarily exposed to at school until grade 3.

Frequency of language of test spoken at home (*ASBG03*) is a background variable from the student questionnaire, asking the students how often the language of test is spoken at home. Students were offered three response categories (i.e., 'I always or almost always speak <language of test> at home', 'I sometimes speak <language of test> and sometimes speak another language at home', and 'I never speak <language of test> at home') (IEA's Progress in Reading Literacy Study, n.d.). The variable showed about 16% of missing information.[56]

Categories of both variables have been combined into two groups. These are, first, students who took the prePIRLS assessment in a native African language and who indicated that their primary language spoken at home is always or sometimes the language of test (57.8% in group Native); and second, students who took the reading assessment in English or Afrikaans and are primarily exposed to these languages at home as well (19.3% in group NonNative). For the purpose of completeness, students whose language of test did not coincide with the home language were summarized in an additional group labeled *Others*. They make 6.7% of the sample.[57] The distribution shows that the share of students primarily exposed to a native African language both at home and at school (group Native) exceeds the group NonNative by far. This is in line with official statistics reflecting that native African languages dominate the countries linguistic landscape (sec. 2.2.1).

8.2.5. Missing Data

Missing data is a common phenomenon in empirical studies, and reasons for missing data to occur are multifaceted. Besides intended missingness, that is information being missing by design, respondents might also refuse to answer, or give answers that are invalid due

[56] The option of using information from the home background questionnaire administered to parents has been discarded, given the large shares of missing information in variables on the frequency of language used by the parents in the home when talking to the child (*ASBH16A*, *ASBH16B*), each reaching up to almost 50% of missingness (own calculations, based on IEA's Progress in Reading Literacy Study, n.d.).

[57] The group Others is not considered for the analysis of this study.

to illegibility (R. Little & Rubin, 2020; Lüdtke et al., 2007). Graham et al. (2003) summarized three issues that may occur in association with missing data. First, reduced sample sizes due to large amounts of missing data might negatively affect the power of parameter estimation. Second, empirical operations often require complete or near complete information matrices because they are not designed to work reliably based on larger shares of missing information. And third, systematic differences between valid and missing data might bias estimation of parameters (Graham, 2003). Accordingly, patterns of missing data need to be investigated and addressed using suitable statistical procedures in order to ensure a precise estimation of standard errors, and hence, allow for a reliable interpretation of results. To that end, this section briefly presents the premises of missing data, and introduces the procedure of Missing-Data-Diagnosis according to Wirtz (2004) in order to explore patterns of missingness.[58] Based thereon, missing data in the South African data set of prePIRLS 2011 is explored, though it must be noted that this only concerns items introduced to the SEM. Lastly, implications for this study are derived and summarized.

Types of Missing Data

Rubin (1976) systematically distinguished between three types of missing data, namely data missing completely at random (MCAR), data missing at random (MAR), and data missing not at random (MNAR) (Rubin, 1976). If data is MCAR, missing values are considered a random sample of the data. The pattern of missingness does neither relate systematically to characteristics of the concerned variable, nor to characteristics of other variables in the data set. If data is MAR, the opposite is true, and patterns of missingness relate systematically to characteristics of other variables in the data set, though, there is no interrelation with the characteristics of the concerned variable. Therefore, data MCAR is considered to be a special case of data MAR. Lastly, if data is MNAR, none of the above assumptions hold true (Lüdtke et al., 2007).

[58] The SPSS syntax applied for data preparation and data analysis is available in the extended appendix (sec. 'Data Preparation (SPSS Syntax)' and 'Missing-Data-Diagnosis (SPSS Syntax)').

Aimed at depicting the different levels of missing data, Lüdtke et al. (2007) used the example of income and age. If data is MCAR, individuals with missing information on income are considered random. That is, missing information do not relate to different levels of income, age or any other variable in the data set. If data is MAR, missing information on income are independent from different levels of income, even when controlling for additional variables, such as age. If data is MNAR, missing information on income are associated with different levels of income, even though other variables such as age have been controlled for (Lüdtke et al., 2007). Hence, Rubin's classification provided an important tool to narrow assumptions regarding the patterns of missing data. Although established in the seventies, it is therefore still taken as the basis for many contemporary approaches towards handling missing data.

Systematization of Missing Data in prePIRLS

Due to its strict assumptions about patterns of missingness, data MCAR is rarely justifiable. However, in international large-scale assessments such as TIMSS and PIRLS, data on student achievement is comprehensibly considered to be MCAR. This is due to the applied multi-matrix sampling design (sec. 8.1.2), where the rotation of booklets for the purpose of reducing testing time results in a random assignment of items to students. Thus, items that have not been randomly assigned by design are assumed to be missing completely at random. However, this assumption only holds true for occurring missingness in items administered in the assessment to measure performance. The logic does not apply to missingness occurring in items of the background questionnaires (Lüdtke et al., 2007).

In prePIRLS 2011, there are three different kinds of missing data in items administered in the contextual background questionnaires. Codes for missing data in prePIRLS 2011 are *omitted or invalid*, *not administered*, and *not applicable*. The code *omitted or invalid* is assigned if an item was left blank or an invalid answer was provided, that is the answer was unreadable or the item was handled falsely. The code *not administered* is used to label questionnaires that were not completed at all, or if a specific item was not included

in the questionnaire. The latter includes items that were left out of the instrument intentionally, items that were mistranslated on national level and, therefore, incomparable internationally, and items that were misprinted. The code *not administered* is thereby used to distinguish missing data from missingness due to non-response. This type of missingness is also referred to as *system missing*. Lastly, the code *not applicable* is used for items linked to filter questions (Foy & Drucker, 2013b).

Missing-Data-Diagnosis

The question remaining is how to detect whether data is missing at random (MAR). An option to do so is the *MCAR-test* after Little (1988). However, the test only proves whether the condition of data being missing completely at random (MCAR) is violated (Wirtz, 2004, p. 112). Thus, whether data is MCAR, or missing not at random (MNAR), cannot be identified empirically using the test (Lüdtke et al., 2007; Wirtz, 2004). Hence, Wirtz (2004) argued that the decision on the nature of missingness needs to be based on profound theoretical knowledge on the subject, paired with statistical testing (Wirtz, 2004). To that end, Wirtz (2004) proposed to investigate the nature of missingness through a two-step approach, referred to as Missing-Data-Diagnosis.

The overall goal of Missing-Data-Diagnosis is to explore what determines missingness, and whether patterns of missingness occur randomly or systematically, that is, MAR or MNAR (Wirtz, 2004). To that end, Wirtz (2004) proposed to check for the following characteristics of missingness in two sequential steps. First, the rate of missingness in each item of interest should be checked using descriptive statistics. That is, to identify cases or groups that systematically omitted the answer to a specific question. As a rule of thumb, Wirtz (2004) posited to investigate missing data when rates of missing information exceed 5% in a single variable, and to exclude a variable from the analysis if rates of missingness exceed 30% (Wirtz, 2004). Additionally, attention should be paid to patterns of missingness, that is whether missing information occur in a group of variables (Wirtz, 2004). Second, missing information in one variable should be examined with the values of

other variables, aiming to identify whether patterns of missingness occur systematically or in relation to other items. To that end, a dummy variable (x), where 0 represents missing information and 1 represents valid information, is associated with an outcome variable (y). Differences in the expression of x and y are compared using a t-test. If differences are significant, it must be assumed that data is missing systematically, that is in association with the outcome variable (y). Lastly, checking for such characteristics, one gains a substantial and valuable overview of the data and patterns of missingness, supporting the choice of an appropriate procedure to address missing data (Wirtz, 2004).

With Missing-Data-Diagnosis, Wirtz (2004) introduced a rather basic approach to the exploration of the nature of missingness. However, many studies refrain from investigating patterns of missing data altogether, simply declaring thereof as MAR or MCAR, given that most empirical models and methods require that. Presumably, this is because the investigation and treatment of data being not missing at random (MNAR) requires sound knowledge in the field of advanced imputation methodologies (Lüdtke et al., 2007), and thus, often exceeds the expertise and resources of research projects and personnel. Hence, Missing-Data-Diagnosis according to Wirtz (2004) offers an easy to apply procedure to start exploring missing data, and functions as a compromise between dismissing the investigation of missingness completely, and accumulating required resources for advanced missing data analyses.
Given that variables introduced to the SEM of this study partly show large amounts of missing information, close to or even above 30% (tab. 2), Missing-Data-Diagnosis according to Wirtz (2004) is performed to explore the nature of the missing data, and make an informed decision on approaching it statistically.

Step 1: Exploring Rates of Missingness

Items utilized for the SEM of this study show two types of missing data. These are systematically missing values, which are missing by design, and thus, are not administered, and missing values due to information being omitted or invalid (Foy et al., 2012). Table

Domain & Subdomain	Variable	Variable Description	N_{Valid}	% Missing		
				Omitted or Invalid	*Not Administered*	*Total % Missing*
Parental Motivation:						
School Invitations for Involvement	ASBH10A (b, r)	Parents feel included in child's education	11571	11.5	15.0	26.5
Educational Aspirations	ASBH18 (c)	Parental expectations of child's educational level	10435	18.8	15.0	33.7
Student Outcomes in Learning:						
Reading Achievement	ASRREA01-05 (*)	Overall reading achievement	15744	n.a.	n.a.	0
Academic Self-Efficacy	ASBDSCR (d)	Students feel confident in reading	13393	14.6	0.3	14.9
Intrinsic Motivation	ASBDSLR (e)	Students like reading	14734	6.1	0.3	6.4
Parental Involvement:						
Supervision	ASBH09A (a, r)	Parents discuss schoolwork	11665	10.9	15.0	25.9
	ASBH09B (a, r)	Parents help with homework	11475	12.1	15.0	27.1
	ASBH09C (a, r)	Parents ensure child sets aside time for homework	11154	14.2	15.0	29.2
	ASBH09D (a, r)	Parents ask what child learned	11280	13.4	15.0	28.4
	ASBH09E (a, r)	Parents check child's homework	11222	13.7	15.0	28.7
Practice	ASBH09F (a, r)	Parents help practice reading	11190	13.9	15.0	28.9
	ASBH09G (a, r)	Parents help practice math	11194	13.9	15.0	28.9
	ASBH09H (a, r)	Parents talk with child about what child reads	11256	13.5	15.0	28.5

Legend:
(*) Scale, treated as observed variable. Details on methods and procedures for scales and plausible values can be found at Joncas & Foy, 2013.
(r) Reverse coding has been applied to ensure concordant pairing of covariance matrices.
(a) Categories are: 1 'Never or almost never', 2 'Once or twice a month', 3 'Once or twice a week', 4 'Every day or almost every day'
(b) Categories are: 1 'Disagree a lot', 2 'Disagree a little', 3 'Agree a little', 4 'Agree a lot'
(c) Categories are: 1 'Grade 9/Standard 7', 2 'Grade 12/Standard 10', 3 'Post secondary training (Vocational training; e.g., College)', 4 'Technikon Diploma', 5 'First degree', 6 'Honours degree/Masters or PhD degree'
(d) Categories are: 1 'Confident', 2 'Somewhat confident', 3 'Not confident'
(e) Categories are: 1 'Like Reading', 2 'Somewhat like reading', 3 'Do not like reading'
(n.a.) Information is not meaningful.

Table 2: Rates of missingness. Unweighted results in percent. Own illustration, based on own analysis using the South African data set of prePIRLS 2011.

2 shows the distribution of percent missingness across all items considered in the models of this study. Furthermore, table 13 in the appendix shows percent missingness in each item by groups of this study (sec. 8.2.4), distinguishing the total percent missingness, and percent of missingness due to items being not administered.[59]

First, table 2 summarizes missingness in items across the complete data set. Here, larger and equally distributed shares of items not administered (15%) are found among most items. An exception are the scales related to student outcomes in learning, where 0.3% of information in the scales of students like reading (*ASDGSLR*) and students feel confident in reading (*ASDGSCR*) have not been administered. Though, according to Wirtz (2004), missing information below 5% do not need to be investigated. Items being omitted or

[59] For the listed analyses, the index variables of students like reading (*ASDGSLR*) and students feel confident in reading (*ASDGSCR*) are used, given that these distinguish missings by items not administered and items omitted or invalid. Descriptive results were computed with SPSS rather than the IDB Analyzer, given that SPSS provides separate counts of items not administered and items omitted or invalid. Thus, descriptive results are unweighted.

invalid range between 6% in students' intrinsic motivation and 19% in parental educational aspirations. Hence, none of the items of interest exceed the threshold of 30%, and thus, remain in the pool of items utilized for this study. Furthermore, table 2 shows that in the battery of items on parental involvement, rates of missing are similarly high, indicating patterns of missingness.

Second, the distribution of missingness by groups (tab. 13 in the appendix) shows that largest shares of missing information occur when there are also no information on either of the grouping variables. For example, 72% of parents, who did not provide information on home resources for learning also omitted a response on the aspiration towards their children's education. Furthermore, shares of items not administered are again equally distributed (e.g., 36% of items on parental involvement activities in group MHRL).

Summarizing the descriptive statistics on rates of missingness, a closer look at items being not administered and items being omitted or invalid is recommended. An exception are the index variables of the scales on students like reading (*ASDGSLR*), and students feel confident in reading (*ASDGSCR*), where items being not administered remain below 5%, regardless of the group.

Step 2: Systematic Associations of Missingness

According to Wirtz (2004), missing information in one variable accompanied by missing information in another hint towards data being missing systematically. On that note, missingness due to items being not administered and missingness due to items being omitted or invalid is further explored. To do so, each item of interest is recoded to 1 when it has been classified as not administered, and to 0 when it has been classified as omitted or invalid, or to 2 when it provides valid information. For the outcome variable, student overall reading achievement (*ASRREA01 - 05*) is used, given that the scale offers a complete data matrix with no missing information. The basis for the evaluation of patterns of missingness is the average difference in student overall reading achievement when associated with missing and valid information in the variables of interest. A significant t-test entails that missingness occurs systematically, and hence, not at random.

Aiming to answer whether respective items not administered are classified as MAR or MNAR, content knowledge on the field of research needs to be acknowledged (Wirtz, 2004). To that end, the *Center for Evaluation & Assessment* (CEA) located at the University of Pretoria in South Africa and responsible for the national assessment of PIRLS and prePIRLS has been contacted on the matter of rates of missingness from items not administered being equally distributed (tab. 2). The CEA was able to confirm that, in the case of prePIRLS 2011, several background questionnaires had been misprinted in the course of the administration, thus leading to an equal distribution of missingness across all items. However, given that misprinted background questionnaires have been randomly distributed among participating schools, and parents of students have been sampled within schools, it is assumed that data not administered is justifiably categorized as MAR.
The assumption of items not administered being MAR is furthermore supported by the results of associating missing information in the variables of interest with an outcome variable, using a t-test. As indicated above, there is no indication for Missing-Data-Diagnosis in the index variables of students like reading (*ASDGSLR*) and students feel confident in reading (*ASDGSCR*) given that information not administered do not exceed the threshold of 5% (Wirtz, 2004). Although, results of respective analysis are included in the table for completeness, they are grayed out for the given reasons. Table 3 summarizes the relevant coefficients and results of the t-test at the 95% confidence level.
Evidently, none of the performed t-tests show a significant difference between the mean reading achievement of those who have valid information on the items of interest, and those who have missing information categorized as not administered on the items of interest. Thus, the results provide a good approximation to the assumption that information being not administered are justifiably classified as MAR.

Other than the level of missingness on items not administered, the determination of the level of missingness on items omitted or invalid is not classified as easily. Table 4 summarizes the relevant coefficients and results of the t-test at the 95% confidence level.

Domain & Subdomain	Variable	Variable Description	Average *Reading Achievement* if information are:		T-test	
			Not Administered	*Valid*	*Point Difference*	*Significant at +/-1.96*
Parental Motivation:						
School Invitations for Involvement	ASBH10A (b, r)	Parents feel included in child's education	321.52	330.51	8.99	1.29
Educational Aspirations	ASBH18 (c)	Parental expectations of child's educational level	321.45	325.55	4.09	0.58
Student Outcomes in Learning:						
Academic Self-Efficacy	ASDGSCR (d)	Students feel confident in reading	363.60	328.23	-35.36	-1.23
Intrinsic Motivation	ASDGSLR (e)	Students like reading	363.60	333.34	-30.26	-1.05
Parental Involvement:						
Supervision	ASBH09A (a, r)	Parents discuss schoolwork	321.52	330.05	8.53	1.22
	ASBH09B (a, r)	Parents help with homework	321.52	331.13	9.61	1.37
	ASBH09C (a, r)	Parents ensure child sets aside time for homework	321.52	332.55	11.03	1.57
	ASBH09D (a, r)	Parents ask what child learned	321.52	332.00	10.48	1.49
	ASBH09E (a, r)	Parents check child's homework	321.52	332.43	10.91	1.55
Practice	ASBH09F (a, r)	Parents help practice reading	321.52	332.48	10.96	1.56
	ASBH09G (a, r)	Parents help practice math	321.52	332.36	10.84	1.54
	ASBH09H (a, r)	Parents talk with child about what child reads	321.52	332.09	10.57	1.50

Legend:
(r) Reverse coding has been applied to ensure concordant pairing of covariance matrices.
(a) Categories are: 1 'Never or almost never', 2 'Once or twice a month', 3 'Once or twice a week', 4 'Every day or almost every day'
(b) Categories are: 1 'Disagree a lot', 2 'Disagree a little', 3 'Agree a little', 4 'Agree a lot'
(c) Categories are: 1 'Grade 9/Standard 7', 2 'Grade 12/Standard 10', 3 'Post secondary training (Vocational training; e.g., College)', 4 'Technikon Diploma', 5 'First degree', 6 'Honours degree/Masters or PhD degree'
(d) Categories are: 1 'Confident', 2 'Somewhat confident', 3 'Not confident'
(e) Categories are: 1 'Like Reading', 2 'Somewhat like reading', 3 'Do not like reading'
(n.a.) Information is not meaningful.

Table 3: Results of Missing-Data-Diagnosis for items being not administered, (after Wirtz, 2004). Own illustration, based on own analysis using the South African data set of prePIRLS 2011.

Results show that the average student achievement in reading differs significantly between items being omitted or invalid, and valid items. Thus, information being omitted or invalid must be considered missing systematically, that is MNAR (Wirtz, 2004).

Missing Data Estimation

Subsequent to the investigation of the nature of missing data after Wirtz (2004), different mechanisms to address the issue are available.

In line with the disadvantages of missing data, Lüdtke et al. (2007) made an effort to provide an overview on statistical procedures available to address missing data. They distinguish between three approaches, including traditional procedures (e.g., pairwise and listwise deletion), imputation-based procedures (e.g., multiple imputation), and model-based procedures (e.g., maximum likelihood estimation). Whereas traditional procedures are considered to be somewhat inappropriate to address missing data, imputation-based and model-based procedures are more modern ways to approach the issue. In comparison

Domain & Subdomain	Variable	Variable Description	Average *Reading Achievement* if information are:		T-test	
			Omitted or Invalid	*Valid*	*Point Difference*	*Significant at +/-1.96*
Parental Motivation:						
School Invitations for Involvement	ASBH10A (b, r)	Parents feel included in child's education	274.01	330.51	56.50	8.19
Educational Aspirations	ASBH18 (c)	Parental expectations of child's educational level	280.63	325.55	44.92	6.88
Student Outcomes in Learning:						
Academic Self-Efficacy	ASDGSCR (d)	Students feel confident in reading	263.57	333.34	69.77	10.70
Intrinsic Motivation	ASDSLR (e)	Students like reading	251.03	328.23	77.21	10.25
Parental Involvement:						
Supervision	ASBH09A (a, r)	Parents discuss schoolwork	271.36	330.05	58.69	9.20
	ASBH09B (a, r)	Parents help with homework	272.15	331.13	58.98	9.44
	ASBH09C (a, r)	Parents ensure child sets aside time for homework	273.01	332.55	59.54	10.08
	ASBH09D (a, r)	Parents ask what child learned	272.28	332.00	59.72	9.40
	ASBH09E (a, r)	Parents check child's homework	272.40	332.43	60.04	10.04
Practice	ASBH09F (a, r)	Parents help practice reading	272.13	332.48	60.35	10.36
	ASBH09G (a, r)	Parents help practice math	272.53	332.36	59.83	10.11
	ASBH09H (a, r)	Parents talk with child about what child reads	272.45	332.09	59.63	9.98

Legend:
(r) Reverse coding has been applied to ensure concordant pairing of covariance matrices.
(a) Categories are: 1 'Never or almost never', 2 'Once or twice a month', 3 'Once or twice a week', 4 'Every day or almost every day'
(b) Categories are: 1 'Disagree a lot', 2 'Disagree a little', 3 'Agree a little', 4 'Agree a lot'
(c) Categories are: 1 'Grade 9/Standard 7', 2 'Grade 12/Standard 10', 3 'Post secondary training (Vocational training; e.g., College)', 4 'Technikon Diploma', 5 'First degree', 6 'Honours degree/Masters or PhD degree'
(d) Categories are: 1 'Confident', 2 'Somewhat confident', 3 'Not confident'
(e) Categories are: 1 'Like Reading', 2 'Somewhat like reading', 3 'Do not like reading'
(n.a.) Information is not meaningful.

Table 4: Results of Missing-Data-Diagnosis for items being omitted or invalid, (after Wirtz, 2004). Own illustration, based on own analysis using the South African data set of prePIRLS 2011.

to imputation-based procedures, model-based procedures are more efficient, given that they simultaneously incorporate approaching missing data and estimating the model in one step (Lüdtke et al., 2007).

A well-tested method in the field of social sciences is the widely applied model-based estimation procedure *Full Information Maximum Likelihood* (FIML). The basic idea of the FIML method is to estimate the likelihood of model parameters for all cases, that are those with valid and missing information. To do so, FIML draws on valid individual information, and estimates the likelihood of model parameters for all cases based thereon. Consequently, missing information are disengaged from the procedure. Thus, the FIML method does not impute missing information, but rather provides an estimation of population parameters applicable to all cases based on valid information available in the data (Lüdtke et al., 2007, p. 112). Finally, the FIML method is assumed to be the "superior method for dealing with missing data in structural equation models" (Enders & Bandalos,

2001, p. 455).

Technically speaking, FIML operates on the estimation of two models, that is the H_1 and H_0 model. The H_1 model is known as the unrestricted model, providing maximum-likelihood estimates of the mean, variance and covariance parameters based on an Expectation - Maximization (EM) algorithm as discussed by Little & Rubin (1987) (B. O. Muthén, 1998-2004, p. 25). On the contrary, the H_0 model describes the specified model or the latent variable model. Comparing H_1 to H_0, a Chi^2 measure of model fit is provided (B. O. Muthén, 1998-2004). Whereas different types of information matrices are available to the FIML estimation, Savalei (2010) suggested to use the observed matrix as a basis for the information matrix, given that this procedure results in more robust standard errors (Savalei, 2010). Conveniently, this approach is the default option in MPlus.

Simulation studies showed that FIML provides more reliable results than other, more traditional approaches such as listwise or pairwise deletion as well as weighting. That is, the application of FIML leads to rather unbiased estimates and standard errors, also when compared to results of imputation-based methods (e.g., Enders & Bandalos, 2001; Graham, 2003; Larsen, 2011; Lüdtke et al., 2007; Wolgast et al., 2017). Hence, Asparouhov & Muthén (2010) suggested to not use multiple imputation, unless there is a specific reason such as computational limitations when estimating the model using FIML (Asparouhov & Muthén, 2022).

Although preconditions of the FIML method are that data is normally distributed and at least MAR, estimated model parameters with FIML have been shown to be reliable, even when the precondition of data being MAR is moderately violated (Wirtz, 2004). However, the FIML method functions particularly reliable when sample sizes are large enough, and rates of missing information do not exceed 30% in single variables (Wirtz, 2004). Both prerequisites are true for the data utilized in this study.

Summary and Implications for the Study

With regard to items of the South African prePIRLS 2011 data set utilized for the SEM of this study, results of Missing-Data-Diagnosis indicated the following.

Overall, descriptive statistics revealed large shares of missingness among items of interest. Though, further investigation revealed that items not administered can be classified as MAR, and hence, missing unsystematically. This finding is based on background knowledge of the data as provided by the CEA, and on differences in average student overall reading achievement between items of interest being either not administered or validly answered. Differences were insignificant, and thus, information are likely to be missing randomly. On the other hand, items being omitted or invalid occurred somewhat systematically, given that there are significant differences between average student overall reading achievement in association with items being omitted or invalid and validly answered. This is true for all items of interest.

That being said, missing data needs to be addressed properly. To that end, this study opts to apply a model-based procedure, that is the estimation of missing information using the FIML method. Although, items omitted or invalid are assumed MNAR, and thus, a precondition for FIML estimation is violated, the sample size of this study is large enough to still ensure a rather precise estimate of model parameters. Furthermore, rates of missingness for all items of interest are below the established threshold of 30% (Wirtz, 2004), considering the distinction of items not administered and items omitted or invalid. Thus, FIML is expected to function well as an estimation method dealing with missing information in this study.

8.3. Results

The following section presents the results of the quantitative study.[60]

First descriptive statistics on the items of interest included in the SEM are presented. Sec-

[60] The SPSS syntax applied for data preparation is available in the extended appendix (sec. 'Data Preparation (SPSS Syntax)').

ond, preparatory work towards building latent constructs within the SEMs is summarized, including results on invariance testing of measurement and path models, and factor analysis. Lastly, results of multi-group SEMs, build upon prior findings of invariance testing are presented. The basis for both, measurement and path models, are group comparisons. As presented in section 8.2.4, groups of interest are socioeconomic low- and high-performers. They are represented by groups drawn from the home resources for learning scale (i.e., LHRL, HHRL and MHRL) and language groups (i.e., Native and NonNative). Given that this study has a particular interest in the (parental) educational investment behavior of South Africa's socioeconomic low-performers, the group of reference is always the one representing low socioeconomic status (i.e., LHRL and Native), compared to groups representing high socioeconomic status (i.e., HHRL, MHRL and NonNative). In summary, pairs for comparison are the following:

- LHRL - HHRL
- LHRL - MHRL
- Native - NonNative

The comparability of respective pairs is verified through invariance testing of measurement and path models. Additional comparisons between pairs, for example between LHRL and NonNative, are not addressed in this study.

8.3.1. Descriptive Analyses

Overall Descriptive Item Statistics

The set of variables utilized for this study (sec. 8.2.3), as well as their statistical properties are summarized in table 5. Aimed at providing a comprehensive overview of the data and data distribution, the table includes information on the domain as well as sub-domain, corresponding (sets of) variables, the number of valid cases (n_{Valid}), the median ($\widetilde{x}$), the mean ($\bar{x}$), and the standard deviation (S.D.). Additionally, a legend is provided for further clarification on relevant information, for example recoding and response categories.

Domain & Subdomain	Variable	Variable Description	N_{Valid}	Median	Mean	S.D.
Parental Motivation:						
School Invitations for Involvement	ASBH10A (b, r)	Parents feel included in child's education	11571	4	3.71	0.65
Educational Aspirations	ASBH18 (c)	Parental expectations of child's educational level	10435	6	4.65	1.64
Student Outcomes in Learning:						
Reading Achievement	ASRREA01-05 (*)	Overall reading achievement	15744	(n.a.)	323.34	112.49
Academic Self-Efficacy	ASBGSCR (*)	Students feel confident in reading	13393	9	9.13	1.74
Intrinsic Motivation	ASBGSLR (*)	Students like reading	14734	10	9.72	1.52
Parental Involvement:						
Supervision	ASBH09A (a, r)	Parents discuss schoolwork	11665	4	3.52	0.72
	ASBH09B (a, r)	Parents help with homework	11475	4	3.55	0.72
	ASBH09C (a, r)	Parents ensure child sets aside time for homework	11154	4	3.49	0.84
	ASBH09D (a, r)	Parents ask what child learned	11280	4	3.59	0.73
	ASBH09E (a, r)	Parents check child's homework	11222	4	3.6	0.73
Practice	ASBH09F (a, r)	Parents help practice reading	11190	4	3.43	0.79
	ASBH09G (a, r)	Parents help practice math	11194	4	3.32	0.84
	ASBH09H (a, r)	Parents talk with child about what child reads	11256	4	3.40	0.81

Legend:
(*) Scale, treated as observed variable. Details on methods and procedures for scales and plausible values can be found at Joncas & Foy, 2013.
(r) Reverse coding has been applied to ensure concordant pairing of covariance matrices.
(a) Categories are: 1 'Never or almost never', 2 'Once or twice a month', 3 'Once or twice a week', 4 'Every day or almost every day'
(b) Categories are: 1 'Disagree a lot', 2 'Disagree a little', 3 'Agree a little', 4 'Agree a lot'
(c) Categories are: 1 'Grade 9/Standard 7', 2 'Grade 12/Standard 10', 3 'Post secondary training (Vocational training; e.g., College)', 4 'Technikon Diploma', 5 'First degree', 6 'Honours degree/Masters or PhD degree'
(n.a.) Information is not meaningful.

Table 5: Descriptive item statistics. Own illustration, based on own analysis using the South African data set of prePIRLS 2011.

Overall, measures of central tendency, here median and mean, show that results do tend to mount at the upper end of an answering scheme. This is notable in all variables, indicating a skewed distribution of the data. For example, on average, parents indicated that they are involved with their children's education every day, regardless of the specific activity ($\widetilde{x} = 4$; $\bar{\text{x}} > 3$). Another notable example is parental educational aspiration (*ASBH18*). Overall, the majority of parents in South Africa tend to aspire for their children to achieve an honors degree, that is a master or doctoral degree ($\widetilde{x} = 6$; $\bar{\text{x}} = 4.65$). Hence, results confirm findings by Howie et al. (2012), who reported that "South African parents have exceptionally high aspirations for their children's education levels and aspire to their undertaking postgraduate education" (Howie et al., 2012, p. XVI). Furthermore, it is also notable that fourth-grade students in South Africa, on average, show upper medium academic self-efficacy ($\bar{\text{x}} = 9.03$), and intrinsic motivation ($\bar{\text{x}} = 9.70$). That is, on a scale with a maximum value of 14 and 15, respectively.[61] Lastly,

[61] It needs to be noted that students were seemingly troubled by reverse-coded items. For example, the scale on student intrinsic motivation to read (*ASBGSLR*) is based on eight items asking about

table 5 shows the average reading score of fourth-grade students in South Africa. That is, proficiency in reading regardless of the language of test ($\bar{x} = 323.34$). Hence, as shown in the international report of PIRLS and prePIRLS 2011, South African students, both in grade 4 and 5, are among the lowest performing students on the international spectrum (I. V. S. Mullis, Martin, Foy, & Drucker, 2012).[62]

Descriptive Item Statistics by Groups

Furthermore, table 6 and table 7 report on the descriptive item statistics, separated by the groups LHRL, HHRL, and MHRL; as well as Native and NonNative.

As for parental motivation, results show that, on average, parents of all groups feel included in their child's education by the school (*ASBH10A*). This is a promising finding, given that the item potentially functions as a motivation for parental involvement activities (Walker et al., 2005). However, less cohesive results are found with regard to educational aspirations (*ASBH18*). In alignment with the hypothesis of status preservation drawn from the framework of rational decision making (sec. 3), parents of group HHRL show highest educational aspirations. On average, they expect their children to achieve a first degree, which is comparable to a Bachelor's degree or alike. However, more surprisingly, parents of both LHRL and MHRL aspire at least a Technikon diploma. That

what one thinks about reading, and how often certain reading activities are carried out outside of school. Whereas the principal component analysis carried out by the IEA showed positive and sufficient factor loadings >.4 for most of the items, negative and notably low factor loadings were observed for reverse coded items. Thus, the reliability coefficient of the scale was estimated to be .46 with only 28% of residual variance explained (Martin & Mullis, 2012c). A rather similar pattern was detected for student academic self-efficacy (*ASBGSCR*). Yet, negative factor loadings came out higher, exceeding .5 for all items (Martin & Mullis, 2012b). Hence, the application of reverse coded items might need to be reconsidered for future assessments in South Africa. Though, acknowledging the measures' limitations, respective scales are still included to the model of this study, given that they are the best available approximations to the constructs of students' personal motivational beliefs.

[62] To ensure comparability of results across educational systems, average scores of reading performance are reported on a scale between 0 - 1000, with a center point of 500, and a standard deviation of 100. The metric of the scale has been established in 2001 and remains stable between assessment cycles (for details see I. V. S. Mullis, Martin, Foy, & Drucker, 2012, pp. 44–45). It is applicable to both PIRLS and prePIRLS 2011. Averages reported here therefore deviate from the results reported in the international report.

			LHRL (25th percentile)		HHRL (75th percentile)		MHRL (Missing)	
Domain & Subdomain	*Variable*	*Variable Description*	*Mean*	*S.D.*	*Mean*	*S.D.*	*Mean*	*S.D.*
Parental Motivation:								
School Invitations for Involvement	ASBH10A (b, r)	Parents feel included in child's education	3.70	0.69	3.72	0.59	3.63	0.76
Educational Aspirations	ASBH18 (c)	Parental expectations of child's educational level	4.43	1.69	5.22	1.26	3.95	1.88
Student Outcomes in Learning:								
Reading Achievement	ASRREA01-05 (*)	Overall reading achievement	309.39	94.37	401.58	131.48	295.38	101.71
Academic Self-Efficacy	ASBGSCR (*)	Students feel confident in reading	9.04	1.63	9.74	2.03	8.88	1.59
Intrinsic Motivation	ASBGSLR (*)	Students like reading	9.68	1.38	10.07	1.76	9.60	1.48
Parental Involvement:								
Supervision	ASBH09A (a, r)	Parents discuss schoolwork	3.41	0.78	3.67	0.59	3.50	0.79
	ASBH09B (a, r)	Parents help with homework	3.46	0.76	3.67	0.61	3.49	0.82
	ASBH09C (a, r)	Parents ensure child sets aside time for homework	3.34	0.95	3.74	0.60	3.35	0.92
	ASBH09D (a, r)	Parents ask what child learned	3.54	0.77	3.69	0.59	3.48	0.87
	ASBH09E (a, r)	Parents check child's homework	3.50	0.81	3.76	0.54	3.50	0.84
Practice	ASBH09F (a, r)	Parents help practice reading	3.33	0.85	3.45	0.74	3.46	0.84
	ASBH09G (a, r)	Parents help practice math	3.23	0.91	3.40	0.73	3.35	0.89
	ASBH09H (a, r)	Parents talk with child about what child reads	3.31	0.88	3.46	0.72	3.38	0.89

Legend:
(*) Scale, treated as observed variable. Details on methods and procedures for scales and plausible values can be found at Joncas & Foy, 2013.
(r) Reverse coding has been applied to ensure concordant pairing of covariance matrices.
(a) Categories are: 1 'Never or almost never', 2 'Once or twice a month', 3 'Once or twice a week', 4 'Every day or almost every day'
(b) Categories are: 1 'Disagree a lot', 2 'Disagree a little', 3 'Agree a little', 4 'Agree a lot'
(c) Categories are: 1 'Grade 9/Standard 7', 2 'Grade 12/Standard 10', 3 'Post secondary training (Vocational training; e.g., College)', 4 'Technikon Diploma', 5 'First degree', 6 'Honours degree/Masters or PhD degree'
(n.a.) Information is not meaningful.

Table 6: Descriptive item statistics by groups of HRL. Own illustration, based on own analysis using the South African data set of prePIRLS 2011.

			Native (African Languages)		NonNative (English/Afrikaans)	
Domain & Subdomain	*Variable*	*Variable Description*	*Mean*	*S.D.*	*Mean*	*S.D.*
Parental Motivation:						
School Invitations for Involvement	ASBH10A (b, r)	Parents feel included in child's education	3.72	0.65	3.69	0.62
Educational Aspirations	ASBH18 (c)	Parental expectations of child's educational level	4.64	1.65	4.70	1.61
Student Outcomes in Learning:						
Reading Achievement	ASRREA01-05 (*)	Overall reading achievement	293.90	88.53	416.62	110.64
Academic Self-Efficacy	ASBGSCR (*)	Students feel confident in reading	9.04	1.65	9.57	1.99
Intrinsic Motivation	ASBGSLR (*)	Students like reading	9.71	1.46	9.94	1.68
Parental Involvement:						
Supervision	ASBH09A (a, r)	Parents discuss schoolwork	3.52	0.71	3.59	0.68
	ASBH09B (a, r)	Parents help with homework	3.56	0.71	3.57	0.69
	ASBH09C (a, r)	Parents ensure child sets aside time for homework	3.39	0.90	3.73	0.62
	ASBH09D (a, r)	Parents ask what child learned	3.58	0.76	3.64	0.65
	ASBH09E (a, r)	Parents check child's homework	3.58	0.74	3.70	0.63
Practice	ASBH09F (a, r)	Parents help practice reading	3.48	0.76	3.37	0.79
	ASBH09G (a, r)	Parents help practice math	3.35	0.84	3.33	0.79
	ASBH09H (a, r)	Parents talk with child about what child reads	3.44	0.80	3.38	0.78

Legend:
(*) Scale, treated as observed variable. Details on methods and procedures for scales and plausible values can be found at Joncas & Foy, 2013.
(r) Reverse coding has been applied to ensure concordant pairing of covariance matrices.
(a) Categories are: 1 'Never or almost never', 2 'Once or twice a month', 3 'Once or twice a week', 4 'Every day or almost every day'
(b) Categories are: 1 'Disagree a lot', 2 'Disagree a little', 3 'Agree a little', 4 'Agree a lot'
(c) Categories are: 1 'Grade 9/Standard 7', 2 'Grade 12/Standard 10', 3 'Post secondary training (Vocational training; e.g., College)', 4 'Technikon Diploma', 5 'First degree', 6 'Honours degree/Masters or PhD degree'
(n.a.) Information is not meaningful.

Table 7: Descriptive item statistics by groups of language. Own illustration, based on own analysis using the South African data set of prePIRLS 2011.

is, a tertiary degree with a technical focus from a Technikon (Junge, 2017). The same is true for both language groups, with parents aspiring an average education of a first degree for their children.

Moreover, results on parental involvement show that parents of all groups report similarly high frequencies of involvement in their child's education at home. On average, all parents of all groups reported to engage with their children's education at least once or twice a week, with a tendency towards daily involvement ($\bar{x} > 3.0$ - 3.5) in many tasks. Again, compared to the theoretical framework of this study, this finding somewhat disagrees with the assumption, that parents with lower socioeconomic capacities lean towards less frequent educational investments, and hence, involvement activities, especially in terms of investing time.

Lastly, student outcomes in learning are presented. Relatively similar results between all groups are found for students' academic self-efficacy and intrinsic motivation ($\bar{x} > 9.0$). Hence, in line with the national average (tab. 5), results show an upper medium self-perception in terms of students' personal motivational beliefs among students of all groups. However, small but notable deviations are found with regard to the perception of self-efficacy among students of MHRL ($\bar{x} = 8.88$), and students of HHRL show a slightly higher intrinsic motivation to read ($\bar{x} = 10.07$). Major differences, however, are found between groups with regard to reading achievement. Recall, the national average was estimated at 323 score points. Whereas students of LHRL attain an average of 309 score points on the reading assessment, students of HHRL average nearly 100 score points more (401 score points). Students of MHRL show the lowest outcomes in reading with an average of 295 score points. A similar result is found for students primarily exposed to native African languages (group Native), averaging at 293 score points, whereas students first and foremost exposed to English or Afrikaans (group NonNative) score an average of 416 score points on the reading assessment.

Factor Analyses

As previously stressed, Fan & Chen (2001) as well as Walker et al. (2005) argued for a multidimensional concept of parental involvement. On that note, Walker et al. (2005) applied a five-item scale to operationalize a single latent construct of home-based parental involvement within the scope of the revised Model of Parental Process (sec. 8.2.3). The construct was well fitting, with a Cronbach's alpha of .82 (Walker et al., 2005). Attempting to mirror the scale of Walker et al. (2005) using suitable items from prePIRLS 2011, exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) are applied.

First, EFA is computed.[63] Comparing a single-factor to a two-factor solution, the Chi^2 test statistic is consulted first. Results of the Satorra-Bentler scaled Chi^2 Difference Test indicate a significant difference between the models ($\Delta Chi^2 = 153.495$; $df = 7$; $p < .001$), implying that one fits better than the other. That being said, fit indices show slightly better model fit for a two-factor solution (CFI/TLI = .991/.980; RMSEA = .023; SRMR = .015).

Whereas a single-factor solution implies a conglomerate of parental involvement activities and is aligned with the proposal of Walker et al. (2005), a two-factor model allows for a more differentiated operationalization of home-based parental involvement activities. Whereas the first factor addresses parental supervision, summarizing items on the frequency that parents talk about the school day and discuss school work as well as supervise homework; the second factor addresses parental activities on practicing mathematics and reading. Furthermore, advantages of a two-factor solution are more precise observations of the effects of home-based parental involvement activities, and hence, more specific evaluations of results.

[63] Exploratory factor analysis was applied using the default settings in Mplus for models including both categorical and continuous variables among others. Hence, MLR was chosen as an estimator, and the rotation method was oblique. Results are based on one plausible value only, given that exploratory factor analysis is not available for imputed data sets. Additionally, the minimum and maximum of extracted factors has been set to 1 and 3, addressing potential issues occurring from the limited set of variables used for the EFA (Kleinke et al., 2017). The Mplus syntax for the computation of the EFA is included in the extended appendix (sec. 'Explanatory Factor Analysis (Mplus Syntax)').

Second, CFA is applied in order to strengthen the assumption of a two-factor solution.[64] Again, comparing a single-factor and a two-factor solution, the Chi^2 test statistic is consulted first. Results of the Satorra-Bentler scaled Chi^2 Difference Test show significant differences in model fit ($\Delta Chi^2 = 102.550$; df = 1; $p < .001$).
Model fit indices of a two-factor solution indicate a well-fitting model (CFI/TLI = .989/.984; RMSEA = .021; SRMR = .018).[65] Hence, the two-factor solution is favored. Table 14 in the appendix shows the standardized factor loadings obtained from CFA.

Based on the results of EFA and CFA, this study opts for a two-factor solution of home-based parental involvement, separating activities into two dimensions, which are *Supervision* and *Practice*.

8.3.2. Invariance Testing

In order to ensure comparability of measurement and path models between groups considered in this study, multi-group invariance testing needs to be carried out. Measurement models in this study are latent constructs of parental supervision of children's education

[64] As in EFA, the default settings of Mplus were mainly kept for CFA as well, including that the factor loading of the first item is fixed to 1, providing a metric to the latent construct. For estimation, MLR has been applied. To obtain the scaling correction factor needed for the Satorra-Bentler scaled Chi^2 Difference Test, only one plausible value has been used to avoid data imputation. For details on the default settings of Mplus see the Mplus User Guide (L. K. Muthén & Muthén, 1998-2017). Furthermore, the Mplus syntax for the computation of the CFA is included in the extended appendix (sec. 'Confirmatory Factor Analysis (Mplus Syntax)').

[65] Although commonly reported, Cronbach's Coefficient Alpha is not considered here. The reasoning is twofold. First, research on reliability measures stemming from the field of psychometrics is rather indecisive whether the Cronbach's Alpha is a suitable measure for the evaluation of scale reliability. For example, Raykov (1997) and Raykov & Marcoulides (2019) provided substantial work, arguing towards the use of Cronbach's Alpha (Raykov, 1997; Raykov & Marcoulides, 2019); whereas Hayes & Coutts (2020) questioned the reliability of the coefficient, given that Cronbachs's Alpha is rather sensitive to sample size and number of items (Hayes & Coutts, 2020). Given that the latent constructs on parental involvement consist of five items in *Supervision*, and three items in *Practice*, the reliability of Cronbach's Alpha might be biased due to the small number of items introduced to the latent constructs (Hayes & Coutts, 2020). Along those lines and because Mplus does not provide the Cronbach's Coefficient Alpha by default for the same reasons (B. O. Muthén, 2008), the coefficient is not reported in this study.

(i.e., *Supervision*), and parental activities of practicing subject related tasks with the child (i.e., *Practice*), as identified in EFA and CFA.

Procedures of invariance testing as well as the criteria to evaluate thereof are applied according to the procedure introduced in section 8.2.2. Invariance testing of measurement models is performed on one data set to avoid imputation, and thus, enable the output of modification indices as well as the performance of the Satorra-Bentler scaled Chi^2 Difference Test (sec. 8.2.1).

First, results of measurement invariance testing are presented, followed by the results of path model invariance, ensuring the equity, and thus, comparability of SEMs between groups.[66]

Measurement invariance

To check for the invariance of measurement models, multi-group CFA is applied, given that this is a common procedure for establishing measurement invariance, especially in cross-cultural contexts (Svetina et al., 2020). The approach is appropriate for the purpose of this study, given that socioeconomically different groups of the South African population are compared. Multi-group CFA is performed prior to building the SEM, aiming to ensure comparability of latent constructs on parental involvement.

This section is structured by levels of invariance, presenting and discussing results for all groups considered in this study. Tables 8 and 9 summarize relevant fit indices as well as differences between them, gradually building levels of configural, metric and scalar measurement invariance. Results of partial metric and partial scalar measurement invariance are considered as well.

First, the measurement model is tested for each group separately. This is called the configural model. Results of the configural models for each group are printed at the top of

[66] The Mplus syntax applied for (partial) measurement invariance testing is available in the extended appendix (sec. 'Measurement Invariance Testing (Mplus Syntax)'). Furthermore, the Mplus syntax applied for testing invariance of path models is available in the extended appendix (sec. 'Structural Equation Model (Mplus Syntax)').

Results of Measurment Invariance for HRL Groups (LHRL-HHRL; LHRL-MHRL)

	Fit Indices				*Satorra-Bentler Scaled Chi² Difference Test*					
	CFI	**ΔCFI**	**RMSEA**	**ΔRMSEA**	**X^2 (T_x)**	**df (d_x)**	**Scaling Correction Factor (c_x)**	**ΔX^2 (TRd)**	**Δdf**	**p**
Configural Model LHRL	0.986		0.034		80.144	19	3.4199			
Configural Model HHRL	0.963		0.036		80.791	19	3.4314			
Configural Model MHRL	0.982		0.025		54.729	19	3.6257			
Baseline Model (LHRL and HHRL)	**0.977**		**0.035**		**160.937**	**38**	**3.4257**			
Metric Invariance	0.974	-0.003	0.035	0.000	184.858	44	3.4117	23.881	6	0.005
Partial Metric Invariance (ASBH09B*, ASBH09H*)	0.976	-0.001	0.034	-0.001	168.858	42	3.4091	7.484	4	0.112
Scalar Invariance	0.966	-0.011	0.037	0.002	231.526	50	3.2831	73.741	12	0.000
Partial Scalar Invariance (ASBH09C*, ASBH09D*)	0.973	-0.004	0.034	-0.001	193.570	48	3.3528	31.758	10	0.000
Baseline Model (LHRL and MHRL)	**0.983**		**0.029**		**134.131**	**38**	**3.5228**			
Metric Invariance	0.981	-0.002	0.029	0.000	157.398	44	3.4882	23.407	6	0.006
Partial Metric Invariance (ASBH09D*)	0.982	-0.001	0.028	-0.001	145.116	43	3.5088	10.777	5	0.056
Scalar Invariance	0.978	-0.005	0.030	0.001	179.988	50	3.3764	46.414	12	0.000
Partial Scalar Invariance (ASBH09A*, ASBH09D*)	0.980	-0.003	0.029	0.000	165.067	48	3.4007	30.247	10	0.001

(*) Items freed along the lines of partial measurment invariance.

Table 8: Summary of results on measurement invariance for pairs of groups by HRL (LHRL-HHRL; LHRL-MHRL) for the measurement models of parental involvement. Own illustration, based on own analysis using the South African data set of prePIRLS 2011.

Results of Measurment Invariance for Language Groups (Native-NonNative)

	Fit Indices				*Satorra-Bentler Scaled Chi² Difference Test*					
	CFI	**ΔCFI**	**RMSEA**	**ΔRMSEA**	**X^2 (T_x)**	**df (d_x)**	**Scaling Correction Factor (c_x)**	**ΔX^2 (TRd)**	**Δdf**	**p**
Configural Model Native	0.993		0.013		41.773	19	5.3555			
Configural Model NonNative	0.986		0.041		91.952	19	2.5341			
Baseline Model (Native and NonNative)	**0.990**		**0.021**		**115.780**	**38**	**3.9448**			
Metric Invariance	0.989	-0.001	0.021	0.000	132.885	44	3.9222	17.060	6	0.009
Partial Metric Invariance (ASBH09D*)	0.990	0.000	0.020	-0.001	124.764	43	3.9218	8.692	5	0.122
Scalar Invariance	0.973	-0.017	0.030	0.009	258.512	50	3.7453	164.274	12	0.000
Partial Scalar Invariance (ASBH09B*, ASBH09C*)	0.987	-0.003	0.021	0.000	147.698	48	3.8168	32.130	10	0.000

(*) Items freed along the lines of partial measurment invariance.

Table 9: Summary of results on measurement invariance for pairs of groups by language (Native-NonNative) for the measurement model of parental involvement. Own illustration, based on own analysis using the South African data set of prePIRLS 2011.

tables 8 and 9. Although the configural model does not serve as a basis for comparison, it provides valuable information on how the measurement models fit each group individually. Configural models thereby add to the evaluation of the results of the baseline model. Overall, the measurement model fits each group well, exceeding the stipulated cut-offs of CFI (>.95) and RMSEA (<.05) (sec. 8.2.1).

Second, the baseline model is estimated. In the sense of the Step-Up approach as suggested by Brown (2015), the baseline model is an unrestricted model across groups serving as a starting point for the evaluation of model restrictions, and the establishment of measurement invariance. The baseline models in tables 8 and 9 summarize the model fit of the unrestricted model for pairs of groups. Evidently, the baseline model fits well for pairs of HRL (LHHRL-HHRL: CFI = .997, RMSEA = .035; LHHRL-MHRL: CFI = .983, RMSEA = .029) pair of language groups (Native-NonNative: CFI = .990, RMSEA = .021). In addition to consulting common fit indices in order to assess the model fit, Svetina et al. (2020) suggested to consider the contribution of each group to the Chi^2 value, since large deviations between contributions of paired groups might point towards a problematic fit of the model for some groups (Svetina et al., 2020). As shown in table 15 in the appendix,

contributions of LHRL-HHRL are 88.010 and 80.927, contributions of LHRL-MHRL are 77.803 and 56.328, and contributions of Native-NonNative are 56.712 and 59.068. Hence, results for MHRL need to be monitored closely, given the larger deviation in Chi^2 contributions occurring from the comparison to the group LHRL.

Third, metric and scalar invariance is tested, gradually introducing equality constraints to the measurement models. For metric measurement invariance, factor loadings are fixed to be equal between groups, and for scalar measurement invariance, both factor loadings and intercepts are fixed to be equal between groups. Models testing for metric and scalar invariance are compared to the baseline model of each pair, and differences in fit (ΔCFI, ΔRMSEA) are reported in tables 8 and 9. Recall, differences in fit indices should not exceed -.010 in CFI, and .015 in RMSEA (sec. 8.2.2, F. F. Chen, 2007). Finally, thresholds concern the decrease of a fit index, indicated by the negative sign in case of CFI, and the positive increase in RMSEA (Dimitrov, 2010). For both, metric and scalar measurement invariance models, CFI and RMSEA are within the cut-offs (CFI $>$.95; RMSEA $<$.05) for all pairs. Furthermore, differences in CFI and RMSEA (ΔCFI, ΔRMSEA) are well within the discussed thresholds, with exceptions of models testing for scalar invariance for the pairs LHRL-HHRL (ΔCFI = -.011), and Native-NonNative (ΔCFI = -.017).

However, the Satorra-Bentler scaled Chi^2 Difference Test turned out to be significant comparing the baseline model to the metric and scalar measurement invariance model. Hence, measurement models with increasing equality constraints are significantly different, and hence, assumingly unequal to the baseline model (Wang, 2013). To address this issue, the establishment of partial measurement invariance, that is loosening the constraints of certain parameters to achieve comparability of models between groups, is a common procedure (sec. 8.2.2 Byrne et al., 1989; Dimitrov, 2010; Kleinke et al., 2017). The decision of which parameters to free from the restrictions is based on the modification indices provided in the Mplus output.

To establish metric invariance, items on parents helping their children with homework (*ASBH09B*) in the latent factor of *Supervision*, and parents talking with the child about

what the child reads (*ASBH09H*) in the latent factor of *Practice* are freed for the pair of LHRL-HHRL, aimed at ensuring metric invariance. Furthermore, the item on parents asking the child what it has learned in school (*ASBH09D*) in the latent factor of *Supervision* has been freed for the pairs of LHRL-MHRL and Native-NonNative, aimed at ensuring metric invariance. By doing so, a total of 25% of the items are freed for the pair of LHRL-HHRL, and a total of 12.5% of items are freed for the pairs LHRL-MHRL as well as Native-NonNative. Whereas the latter is well within the thresholds of 20%, items freed from restrictions of the pair LHRL-HHRL slightly exceed the threshold (Dimitrov, 2010). Though this unproblematic, as long as this is reported in the results of measurement invariance. Finally, loosening restrictions for the respective items led to insignificant results between pairs of groups, suggesting the successful establishment of partial metric measurement invariance.

Repeating respective steps in order to also establish (partial) scalar invariance remains unsuccessful. Hence, the Satorra-Bentler scaled Chi^2 Difference Test turned out to be significant, even after freeing certain items from restrictions according to modification indices. Consequently, scalar measurement invariance is rejected for the models of this study.

Although (strict) scalar invariance is considered the gold standard for equitable comparisons between groups, (partial) metric invariance has been claimed to be sufficient as well as realistically achievable in applied research (T. D. Little, 1997; Meredith, 1993). To conclude, the results of measurement invariance testing point to the successful establishment of weak factorial invariance, that is metric measurement invariance. Thus, the comparability of factor loadings is ensured.

Invariance of Path Models

In order to also check for the invariance of path models, the restricted SEM is compared to the unrestricted SEM (Kleinke et al., 2017). Unlike the previous analyses on measurement invariance testing, analyses on the invariance of path models test the entire model, and

Results of invariance of path models				
	Differences in fit indices			
	CFI	**ΔCFI**	**RMSEA**	**ΔRMSEA**
LHRL-HHRL				
SEM unrestricted	0.953		0.034	
SEM restricted	0.926	-0.027	0.040	0.006
LHRL-MHRL				
SEM unrestricted	0.969		0.025	
SEM restricted	0.965	-0.004	0.025	0.000
Native-NonNative				
SEM unrestricted	0.969		0.022	
SEM restricted	0.964	-0.005	0.022	0.000

Table 10: Summary of results on invariance of path models of groups by HRL (LHRL-HHRL; LHRL-MHRL), and by language (Native-NonNative), considering partial metric measurement invariance in latent constructs of parental involvement. Own illustration, based on own analysis using the South African data set of prePIRLS 2011.

thus, all five plausible values are considered through imputation. Hence, the modification indices are unavailable, and the Satorra-Bentler scaled Chi^2 Difference Test cannot be performed. Evaluation of the invariance of path models is based on differences in the fit indices CFI and RMSEA (ΔCFI, ΔRMSEA). Table 10 summarizes the results of testing for invariance of path models, based on partial metric measurement invariance.

For the pairs of LHRL-MHRL as well as Native-NonNative, differences in CFI and RMSEA stay within the postulated thresholds (sec. 8.2.2). Only for pair LHRL-HHRL, differences in CFI are exceeding (ΔCFI = -0.027, ΔRMSEA = 0.006). Hence, it can be assumed that structural paths of the SEM estimated for the pair LHRL-HHRL are unequal. Respective findings are not surprising, given that the groups LHRL and HHRL assumingly consist of diverse ethnic, cultural, and socioeconomic characteristics (sec. 2), and thus, show considerably different patterns of educational investment behavior. Consequently, structural paths of the pair LHRL-HHRL are freed from equality constraints and paths established in the SEM cannot be directly compared between the two groups.

8.3.3. Results of Structural Equation Models

In this study, multi-group SEMs are estimated to analyze the association of parental motivation and parental involvement activities, and of parental involvement activities

and student outcomes in learning. In alignment with the theoretical framework (sec. 3), linearity between the phases of parental motivation, parental involvement activities, and student outcomes is assumed. Accordingly, indirect effects such as moderation or mediation are untended in this study.

The following section presents the results of the multi-group SEM, separately for each pair considered in this study. The prerequisites thereof, which are the establishment of (partial) metric measurement invariance and the invariance of path models, have been ensured in the preceding section. Results are presented first for the pair LHRL-HHRL, with group LHRL set as the reference group. This model is considered the core model of this study, and thus, an in-depth description of the results is provided. Results for the additional pairs LHRL-MHRL and Native-NonNative each emphasize a particular topic. For the pair LHRL-MHRL, attention is paid to the question of what the results imply with regard to the composition of the group MHRL. For pair Native-NonNative, the focus is on the question of whether language groups are suitable proxies for the socioeconomic status of the South African population.

As pointed out in section 8.3.2, comparisons are based on the unrestricted path model for the pair LHRL-HHRL; and on the restricted path models for the pairs LHRL-MHRL as well as Native-NonNative. Consequently, estimated coefficients for the group LHRL are not congruent between groups of HRL. Furthermore, comparability across pairs has not been verified through invariance analysis, and thus, comparisons of groups across pairs is not recommended. Results of multi-group SEMs are presented in figure 14 for the pair LHRL-HHRL, figure 15 for the pair LHRL-MHRL, and figure 16 for the pair Native-NonNative.[67]

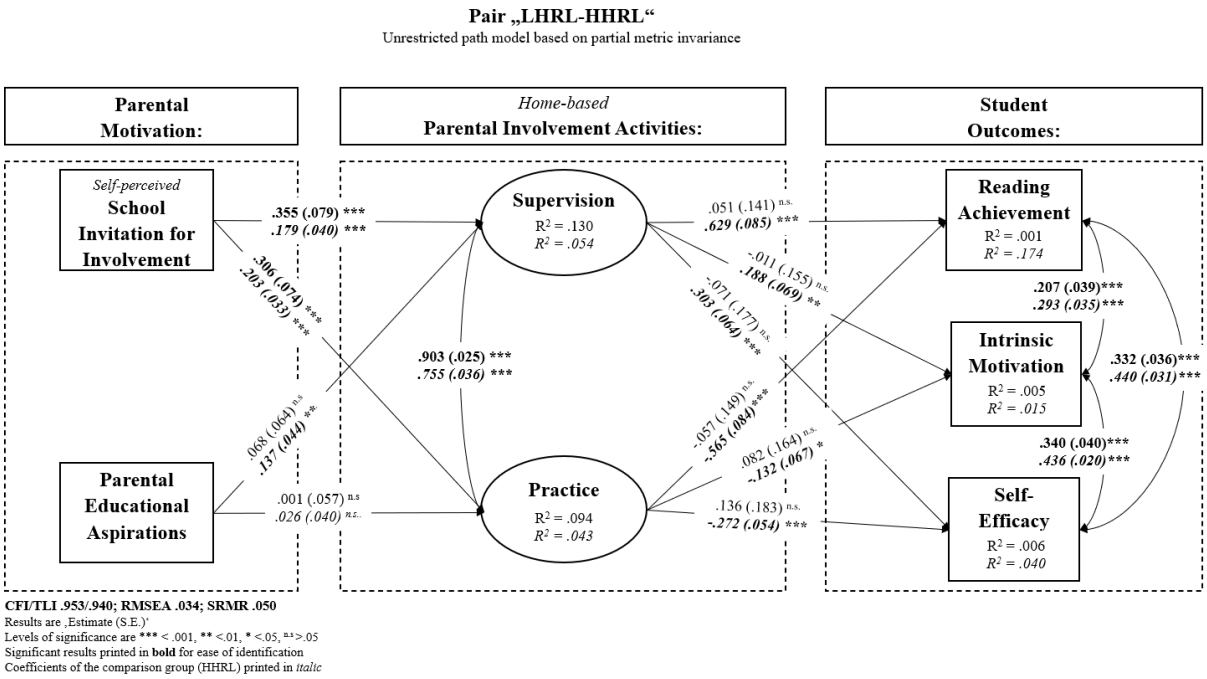

Figure 14: Results of the *unrestricted* path model for the pair LHRL-HHRL, accounting for partial metric invariance. Own illustration, based on own analysis using the South African data set of prePIRLS 2011.

Pair LHRL-HHRL

Overall, the unrestricted SEM of the pair LHRL-HHRL presented in figure 14 fits well (CFI/TLI = .953/.940; RMSEA = .034; SRMR = .050), with a slight deviation in TLI (<.95). Furthermore, the SEM explains about 13% of variance in *Supervision* and 9.4% of variance in *Practice* for group LHRL, and 5.4% of variance in *Supervision* and 4.3% in *Practice* for group HHRL. In turn, about 90 - 95% of the variance in parental involvement activities is due to information not included in the SEM presented here. Similarly low explanatory power is found with regard to student outcomes in learning, where the model explains very little shares of variance. Though, the only notable effect is found for group

[67] Recall, prePIRLS 2011 is a cross-sectional study, and thus, results cannot be interpreted as causal inferences (sec. 8.1). On that note, the terms *association* or *effect* are merely understood in a statistical sense and are used interchangeably in this section. The Mplus syntax applied for multi-group structural equation modeling is available in the extended appendix (sec. 'Structural Equation Model (Mplus Syntax)').

HHRL, where the model explains 17.4% of student overall reading achievement.

Results show that latent constructs of *Supervision* and *Practice* correlate highly with each other (LHRL: $r = .903$; HHRL: $r = .755$). From an empirical point of view, this is somewhat expected as the items considered for the latent constructs of *Supervision* and *Practice* were collected in a battery of items, aimed at representing parental involvement. Especially for group LHRL, the high correlation coefficient raises the question whether a one-factor solution would have been more suitable, given that constructs do not seem to be as selective. A possible yet content-related reason for the high association between latent constructs, especially in group LHRL, could be that parents of group LHRL consider a high level of involvement socially desirable, and therefore, state to be equally involved in all aspects of parental involvement. On the contrary, the association between *Supervision* and *Practice* for group HHRL is less strong. Hence, parents who supervise their children's schoolwork do not get into the habit of practicing reading and mathematics as regularly as parents of the reference group (LHRL). That is, they are more selective in their actions.
Correlation coefficients for student outcomes show similar patterns between groups, with slightly stronger associations between all outcome variables for group HHRL ($r = .293/ .440/ .436$). Hence, reading achievement of children of higher socioeconomic background is associated with higher students' intrinsic motivation and self-efficacy. The relation is more prominent compared to the reference group (LHRL). Strongest effects for group LHRL are found between intrinsic motivation and self-efficacy ($r = .340$); and for group HHRL between reading achievement and self-efficacy ($r = .440$).

With regard to the association of parental motivation and parental involvement activities, parents' perceptions of invitations for involvement from the school are significantly positively associated with both factors of parental involvement (LHRL: $\beta = .355/.306$; HHRL: $\beta = .179/.203$). However, effects are slightly stronger in the group LHRL, prompting the assumption that parents' perceptions of invitations for involvement from the school have

a more meaningful effect on the educational investment behavior of socioeconomic low-performing parents. Furthermore, parental educational aspirations are not significantly associated with parental involvement in the group LHRL. This finding contrasts the results of descriptive statistics, showing that parents of different socioeconomic states have similarly high aspirations for their children's education (sec. 8.3.1). Hence, educational investments to support their children achieving those aspirations would be assumed. For group HHRL, parental aspirations are significantly associated with parental involvement activities of *Supervision*. The finding substantiates the theoretical assumptions of EV-Theory. That is, socioeconomic high-performing parents and their children experience a higher risk of loss of status (sec. 3.2.2), and thus, tend to invest more in order to secure (the inheritance of) family socioeconomic status, for example, through helping the child to a higher educational degree. Hence, it is plausible to find that educational aspirations motivate actions towards parental involvement in group HHRL, though, the effect is rather small ($\beta = .137$). Interestingly, the effect does not appear in the association of parental educational aspirations and parental activities of *Practice*, indicating that the fear of loss of status among parents of higher socioeconomic status does not lead them to invest more time in practicing a subject with the child (e.g., mathematics or reading).

With regard to the association of parental involvement and student outcomes in learning, the SEM shows insignificant effects of parental involvement in group LHRL, meaning that none of the involvement activities regarding *Supervision* or *Practice* have notable effects on student outcomes in learning, including the development of reading achievement as well as students' personal motivational beliefs. In other words, the quantity of involvement activities of socioeconomic low-performing parents does not improve student outcomes in learning. That is, despite the fact that parents tend to get involved quite regularly in the South African population as well as within groups (sec. 8.3.1). Therefore, the subsequent questions is how to evaluate and possibly improve the quality of parental involvement activities among socioeconomic low-performers.

Contradicting the results for group LHRL, all effects are either significantly positive or

negative for group HHRL, entailing that involvement activities of socioeconomic high-performing parents are more effective than the involvement activities of the reference group (LHRL). This is especially true for parental activities of *Supervision* in group HHRL, which have a significantly positive effect on student overall reading achievement (HHRL: β = .629). However, coefficients turned out negative for all associations of parental activities of *Practice* with student outcomes in learning. Hence, if parents of group HHRL get into the habit of practicing both mathematics and reading with their children, chances are that respective activities of *Practice* will affect student outcomes negatively. Congruent to the effect of *Supervision*, the most notable result in *Practice* is found in association with students' overall reading achievement (β = -.565).

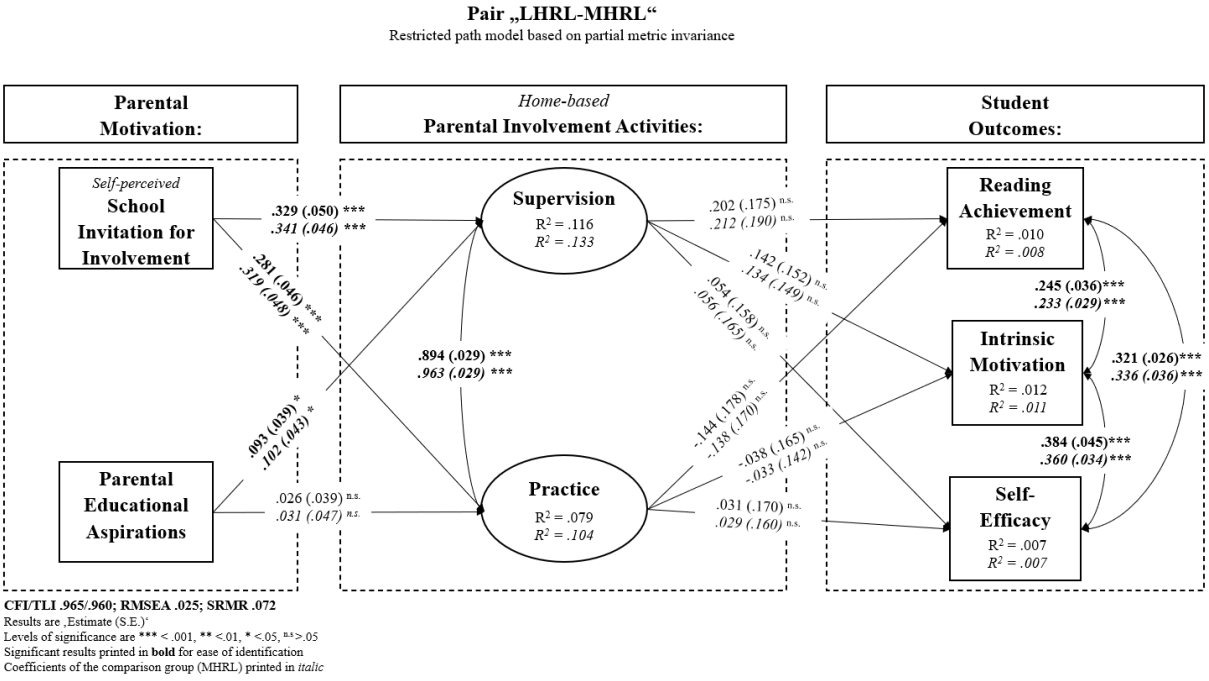

Figure 15: Results of the *restricted* path model for the pair LHRL-MHRL, accounting for partial metric invariance. Own illustration, based on own analysis using the South African data set of prePIRLS 2011.

Pair LHRL-MHRL

For the pair LHRL-MHRL, the multi-group SEM provides information on differences and similarities, in order to learn more about group MHRL. This is because the socioeconomic

status of group MHRL is unclear, given that parents in this group omitted information in that regard. Figure 15 summarizes the results of the restricted multi-group SEM.

Overall, the restricted model of pair LHRL-MHRL fits well (CFI/TLI = .965/.960; RMSEA = .025; SRMR = .072), with a slight deviation in SRMR (>.05).[68] Furthermore, the SEM explains 11.6% of variance in *Supervision* and 7.9% in *Practice* for group LHRL, and 13.3% of variance in *Supervision* and 10.4% in *Practice* for group MHRL. Shares of variance explained with regard to student outcomes in learning are rather small, not exceeding 1% in all of the considered outcomes and for both groups.

Latent constructs of *Supervision* and *Practice* correlate highly with each other (LHRL: $r = .894$; MHRL: $r = .963$). Correlation coefficients between student outcomes are less strong, and show similar patterns between groups, with the strongest associations between students' intrinsic motivation and students' self-efficacy (LHRL: $r = .384$; MHRL: $r = .360$). Furthermore, the correlation between students' self-efficacy and reading achievement (LHRL: $r = .321$; MHRL: $r = .336$) are stronger than correlations between students' intrinsic motivation and reading achievement (LHRL: $r = .245$; MHRL: $r = .233$).

With regard to the association of parental motivation and parental involvement activities, coefficients behave similarly for both groups. Hence, parents' perceptions of invitations for involvement from the school have significantly positive effects on parental involvement activities, that is *Supervision* and *Practice* (LHRL: $\beta = .329/.281$; MHRL: $\beta = .341/.319$). Interestingly, parental educational aspirations show a small but significant effect on *Supervision* (LHRL: $\beta = .093$; MHRL: $\beta = .102$). Therefore, parental activities of *Supervision* are allegedly motivated by parental educational aspirations in both groups of the pair LHRL-MHRL. This finding contradicts the theoretical framework, assuming that motivation of socioeconomic low-performers to invest in education does not overlay the expectations of costs and outcomes (sec. 3.2.2). However, given the size of the coefficients

[68] Recall, fit indices as well as coefficients deviate between the models of pairs LHRL-HHRL and LHRL-MHRL. That is due to differences in restrictions of the path models. The model of pair LHRL-HHRL estimates paths freely, based on results of invariance testing of path models; and the model of pair LHRL-MHRL applies restricted paths.

being rather small, and the fact that there is no solid information on who belongs to group MHRL, these results are not strongly reliable. Finally, the effect of parental educational aspirations is insignificant for involvement activities of *Practice* in both groups.
With regard to the association of parental involvement activities and student outcomes in learning, none of the effects in either of the groups are significant. Apart from descriptive findings on the frequencies of parental involvement activities, respective results are evidence for the circumstance that parental involvement activities of socioeconomic low-performers do not have a meaningful effect on student outcomes in learning. This is true for both groups of the pair LHRL-MHRL.

With regard to the question of whether group MHRL behaves similar to group LHRL, results indicate that this is the case. Assuming that parental involvement is strongly determined by parents' socioeconomic status, similarities of coefficients obtained from the SEM indicate that both groups (i.e. LHRL and MHRL) consist of comparable characteristics, and thus, result in similar patterns of parental educational investment behavior. However, further empirical investigations such as a simulation study would be necessary to substantiate and validate respective assumptions.

Pair Native-NonNative

The pair Native-NonNative focuses on the question whether language groups serve as a suitable proxy for socioeconomic status in the South African population. The underlying assumptions have been derived from official statistics reported in section 2. To that end, observed patterns of pairs LHRL-HHRL and Native-NonNative are put next to each other, though, the comparability between them is not statistically ensured. Figure 16 summarizes the results of the restricted multi-group SEM for pair Native-NonNative.
Overall, the restricted model of Native-NonNative fits well (CFI/TLI = .964/.959; RMSEA = .022; SRMR = .051). The SEM explains 14.6% of variance in *Supervision* and 11.3% of variance in *Practice* in group Native, and 12.8% of variance in *Supervision* and 8.1% of variance in *Practice* in group NonNative. In line with the result of the remaining

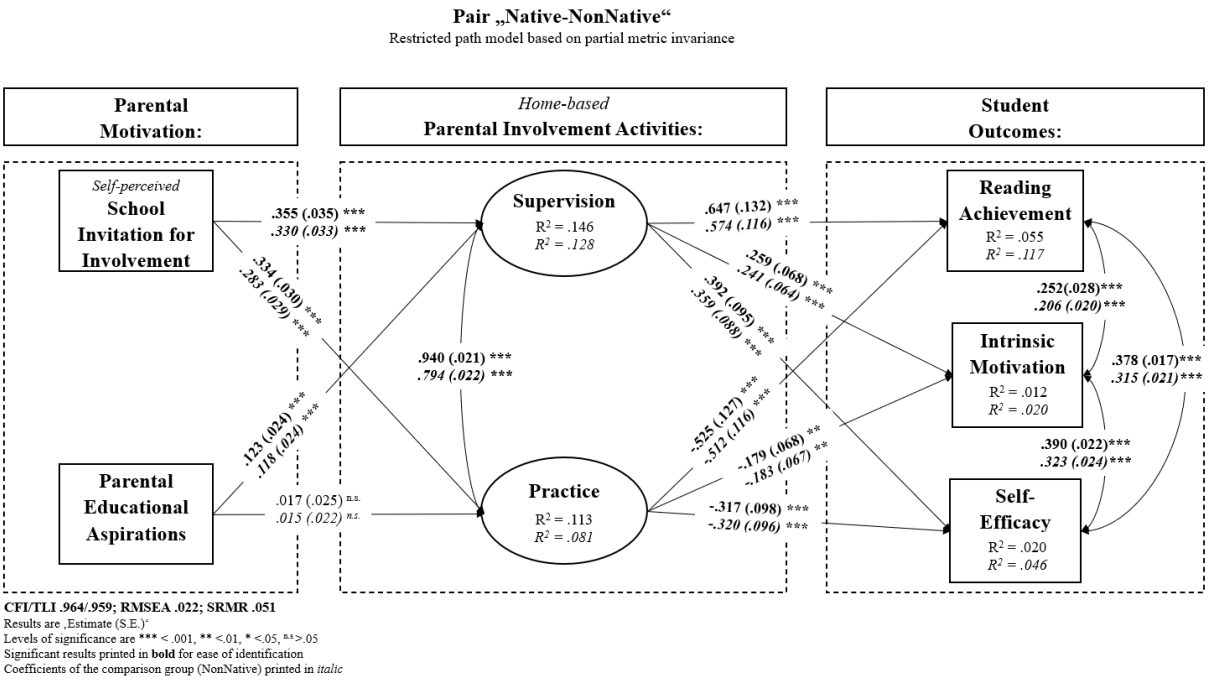

Figure 16: Results of the *restricted* path model for the pair Native-NonNative, accounting for partial metric invariance. Own illustration, based on own analysis using the South African data set of prePIRLS 2011.

pairs, the latent construct of *Supervision* is explained slightly better, though the explanatory power of the SEM altogether is not very high. Shares of variance explained with regard to student outcomes in learning are rather small, with an exception of the effect of student overall reading achievement in group NonNative, where the model explains 11.7% of variance.

Latent constructs of *Supervision* and *Practice* correlate highly with each other (Native: $r = .940$; NonNative: $r = .794$). Evidently, the coefficients are higher for group Native than for group NonNative. These patterns are similar to those found for the pair LHRL-HHRL. The correlation coefficients between student outcomes in learning are less strong, and show similar patterns between groups, with slightly stronger associations between outcomes for group Native ($r = .252/.378/.390$). As observed for pairs LHRL-HHRL and LHRL-MHRL, weakest correlations are found between students' intrinsic motivation and reading achievement (Native: $r = .252$; NonNative: $r = .206$).

With regard to the association of parental motivation and parental involvement activities, the strongest effects are found for group Native. Parents' perceptions of invitations for involvement from the school foster parental involvement activities with regard to *Supervision* and *Practice* (β = .355/.334). Less strong yet still significantly positive effects are obtained for group NonNative (β = .330/.283). As in all other models, self-perceived invitations by the school function as a reliable motive for parents to get involved, though, effects are less strong among socioeconomic high-performing groups. Furthermore, parental educational aspirations have a significantly positive effect on parental supervision (Native: β = .123; NonNative: β = .118). Hence, the higher educational aspirations are, the more often parents get involved in home-based activities of *Supervision*. Similar to pair LHRL-MHRL, effects are rather small. However, contradicting the theoretical assumptions of EV-Theory (sec. 3.2.2), parental educational aspirations are high (sec. 8.3.1), and significantly motivate home-based parental involvement activities of *Supervision* in the reference group (Native).

With regard to the association of parental involvement activities and student outcomes in learning, activities of *Supervision* have a significantly positive effect on learning outcomes in both groups of the pair Native-NonNative. Hence, the more parents supervise their children's school work, the higher is students' overall reading achievement (Native: β = .647; NonNative: β = .574), students' intrinsic motivation (Native: β = .259; NonNative: β = .241) and students' self-efficacy (Native: β = .392; NonNative: β = .359). On the contrary, *Practice* turned out to have a significantly negative effect on student outcomes in both groups. Other than activities of *Supervision*, parental activities of *Practice* significantly impair students' overall reading achievement (Native: β = -.525; NonNative: β = -.512), students' intrinsic motivation (Native: β = -.179; NonNative: β = -.183), and students' self-efficacy (Native: β = -.317; NonNative: β = -.320).

Generally, the effect of parental involvement activities in *Supervision* and *Practice* are comparably strong in both groups of the pair Native-NonNative. Though, significant effects in group Native are surprising, given that this group is considered to represent

low socioeconomic status. Thus, significantly positive associations of parental involvement activities with student outcomes in learning are unexpected. Furthermore, effects of *Supervision* and *Practice* on reading achievement are rather high in group Native (β = .647/-.525), and thus, within the range of group NonNative and HHRL. The respective finding is especially interesting, because average reading achievement of students in group Native equals 294 score points (tab. 7), which is below the average reading achievement of South African fourth-grade students (323 score points, tab. 5) and below the average reading achievement of socioeconomic low-performing groups (LHRL: 309 score points; MHRL: 295 score points, tab. 6). The remaining question is, whether parental involvement deviates in quality between groups, and whether there are additional reasons for student outcomes in learning remaining comparably low.

With regard to the question of whether language groups serve as a suitable proxy for socioeconomic status in South Africa, results of the SEM are contradictory. Although students in group Native have the lowest average students' overall reading achievement among all groups, parental activities of *Supervision* have significantly positive effects on student outcomes in learning; and parental activities of *Practice* have significantly negative effects on student outcomes in learning. A possible explanation for the discrepancy could be that contrasting only two language groups, here Native and NonNative, might insufficiently discriminate patterns of parental involvement activities. A practical solution would be to either account for additional language groups, or to pair items on language of test and home language with home resources for learning, yielding in a more precise grouping variable. The latter was attempted for this study, however, groups became rather small and impractical for multi-group SEM.

8.3.4. Summary

To summarize, results of quantitative analyses showed the following with regard to the theory-driven hypotheses presented in section 5.

Descriptive Results

Results of descriptive analyses showed that parents of all groups reported rather high educational aspirations for their children ($\bar{x} > 4$). Furthermore, parents reported to be satisfied with school efforts to include them in their child's education ($\bar{x} > 4$). With regard to home-based parental educational involvement, parents of all groups reported high frequencies of involvement activities, leaning towards daily involvement. Hence, hypotheses $H_{1.1}$ and $H_{2.1}$ are considered true for all groups. Regarding the average overall reading achievement, South African fourth-grade students differ greatly between groups (tab. 6, tab. 7). Lowest average achievement in reading was observed in group MHRL (295 score points), and in group Native (294 score points). Regardless of differences in overall reading achievement, students of all groups reported similar levels of personal motivational beliefs, that is with regard to self-efficacy and intrinsic motivation towards reading.

Invariance Testing

Invariance testing confirmed (partial) metric measurement invariance for two measurement models of parental involvement (i.e., *Supervision* and *Practice*) for all pairs of groups utilized for this study. Furthermore, invariance of path models was successfully established for all pairs, apart from LHRL-HHRL. Though, this result is considered a reflection of the assumption of South Africa consisting of multiple, fundamentally different societies (Haferburg & Osmanovich, 2017).

Results of multi-group SEM

Results of multi-group SEMs presented a mixed picture in regard to parental educational investment behavior of socioeconomic low-performers in South Africa. Hence, finding distinct answers to the hypotheses of this study (sec. 5) is not straightforward, and the nature of the data, groups, and results call for separate evaluations of each pair.
First, results of the multi-group SEM showed that parents perceptions of school efforts to include them in their child's education are significantly associated with parental home-

based involvement activities between groups, and across pairs. Hence, hypothesis $H_{3.1}$ is confirmed, and the corresponding null hypothesis $H_{3.0}$ is rejected. A more complex pattern of results presents itself for the association between parental educational aspirations and involvement activities. For instance, insignificant associations of parental educational aspirations with involvement activities of *Supervision* and *Practice* are observed for group LHRL in pair LHRL-HHRL, and hence, the null hypothesis $H_{4.0}$ is considered true. Moreover, insignificant results are obtained for the association of parental educational aspirations with activities of *Practice*. Hence, the null hypothesis $H_{4.0}$ is considered true. However, for the remaining groups of socioeconomic low-performers (LHRL, Native) and pairs (LHRL-MHRL, Native-NonNative), significant associations between parental educational aspirations and involvement activities of *Supervision* are obtained, and hence, hypothesis $H_{4.1}$ is considered true for those pairs. Given that respective coefficients are rather small, and response patterns on status of parental educational aspirations are assumed to be affected by social desirability, respective findings must be handled with caution.

Second, results of multi-group SEM showed that involvement activities of *Supervision* are positively associated with student outcomes in learning ($H_{5.1}$). Involvement activities of *Practice* on the other hand resulted in negative effects ($H_{5.0}$), and thus, foster a decrease in student overall reading achievement as well as students' personal motivational beliefs. Lastly, results showed that for group LHRL in pair LHRL-HHRL and for pair LHRL-MHRL, neither *Supervision* nor *Practice* are significantly associated with student outcomes in learning. Hence, the null hypotheses $H_{6.0}$ and $H_{7.0}$ are considered true for these groups and pairs, respectively. However, a particular case is group Native, which behaves similar to its comparison group NonNative, and thus, showed significant associations between parental involvement activities and student outcomes in learning. For this group, hypotheses $H_{6.1}$ and $H_{7.1}$ are considered true. However, results of all models showed that dimensions of parental involvement activities are positively associated with each, as well as student overall reading achievement with students' personal motivational beliefs. Hence, hypotheses $H_{8.1}$ and $H_{9.1}$ are considered true.

9. Discussion of Results

The following section embeds the results obtained from the qualitative and the quantitative study within the current state of research, and evaluates the findings accordingly. Attention is paid to the leading aspects of this study, which is parental involvement activities within the context of educational investment behavior among socioeconomic low-performers in South Africa. Qualitative results are discussed first, followed by quantitative results. As argued above, this study does not meet the criteria of a mixed-methods design, and hence, both studies are discussed separately. However, the complementary character of both studies with regard to the understanding of educational investment behavior of socioeconomic low-performers in South Africa is acknowledged, and findings of both studies are viewed in conjunction, if applicable.

Sequential to the discussion of results, this section concludes with an exposition of advantages and limitations drawn from the empirical results of this study.

9.1. Discussion of Qualitative Results

The qualitative study utilized data obtained from problem-centered interviews with street vendors in Cape Town/South Africa, and resulted in three general types of investment behavior. These are type I *Targeted*, type II *Thwarted*, and type III *Inconsistent*. Interviewees of the qualitative study are considered to be socioeconomically low-performing due to their informal occupation as street vendors, and their level of income, which is below the national poverty line of upper middle income countries (sec. 2.2.3). Given their considerably low standard of living and in line with EV-Theory, street vendors were not expected to be motivated nor to take educational investments.

All identified types show a high motivation for educational investments, but type II and III do not transfer motivation into action. Respective investment behavior can be translated to the terminology of EV-Theory: The motivation for educational investments (U+cSV) results from the value of an investment's utility (U), the expected costs (c), and the risk of loss of status (SV). If the latter is not essential to an actor, which is likely to be the case for individuals of lower socioeconomic status, and hence, street vendors interviewed

in this study, the motivation for educational investments is dissociated from the need for status preservation (cSV=0). Thus, the motivation must ultimately result from the returns to education itself (U). The expected risk of an educational investment is based on the expected educational costs (c), as well as the expected probability for an investment to be successful (p). Hence, the more uncertain the expectation of an educational investment, the higher is the considered risk to take it (sec. 3.2.2). As briefly touched in section 4, the element of time has been identified as a relevant factor for the equation of educational investment behavior as well (G. S. Becker, 1962; Esser, 1999; J. Heckhausen & Heckhausen, 2008; Nerdinger, 1995), and a high socioeconomic status has been identified to be positively associated with the pursuit of more time-consuming educational pathways (Conley, 2001; Light & Strayer, 2000).

For street vendors interviewed in this study, the costs of educational investments are particularly high, considering their occupational and socioeconomic status. This has been expressed by the dilemma earning vs. learning, entailing that street vendors struggle to balance time to generate earnings with time to invest in education. Hence, type II and III show investment behavior in accordance with the summarized theoretical and empirical conceptions.

However, in contradiction to investment behavior according to type II and type III, type I *Targeted* is not only motivated to invest in education, but also takes educational investments. That is despite a presumably precarious socioeconomic status, and lack of time and monetary resources to invest in education. Recall, a change from educational non-investments to educational investments is only possible if the motivation exceeds the risks (Esser, 1999). That is true if the expectation of the probability for success (p), which is presumably small in the case of socioeconomic low-performing street vendors, is overruled by the motivation for an educational investment, and thus, the risks for investments (C/p) are no longer in the way. According to EV-Theory, the investment behavior of type I can only be explained through an unexpectedly high educational motivation among the focus group.

Follow-up studies could explore the investment behavior of type I further, aiming to

discover particular reasons for the motivation for educational investments, potentially against the background of distinct individual living situations. That is, first, to back up the results of this study, and address its limitations (sec. 7.3); and, second, to extend the understanding and disclosure of active educational investment behavior among socioeconomic low-performers in South Africa.

Picking up on the dilemma earning vs. learning as well as resulting non-investments in education, Mulligan (1997) pointed to the importance of parental involvement. He stated that those who invest less time and energy in their children's educational attainment diminish their children's capacities for future states of living, and add to the reproduction of socioeconomic inequality across generations (Attewell & Newman, 2010, p. 8). To put it with Esser (1999): Parents who refrain from educational investments ensure that "the apple doesn't fall far from the tree" (Esser, 1999, p. 265). Along the lines of results on the educational investment behavior of socioeconomic low-performing street vendors reflected in type I, the question is whether similar patterns of investment behavior are obtainable among South African socioeconomically low-performing parents as well. That is, becoming involved in their children's education, and hence, counteracting theoretical conceptions of rational decision making. To explore thereof, the quantitative study was built.

9.2. Discussion of Quantitative Results

Motivated by the results of the qualitative study, sequential quantitative analyses aimed at exploring educational investment further. To that end, the concept of parental involvement was utilized as a form of educational investment behavior. Multi-group SEMs were built to analyze the association between parental educational motivation and parental involvement activities, as well as involvement activities and student outcomes in learning. Furthermore, the quantitative study allowed for contrasting results of socioeconomic low-performers to socioeconomic high-performers, aimed at accounting for the two-tier society of South Africa (sec. 2). For that purpose, the reference groups (LHRL, Native)

were contrasted to the groups of comparison (HHRL, MHRL, NonNative).
The following discussion of results is structured by the two phases derived from the theoretical framework and accounted for in the SEM. Phases are motivational factors of parental involvement activities, and parental involvement activities with regard to student outcomes in learning. Findings of this study are furthermore situated with regard to the current state of research on parental involvement within and beyond the South African context.

Motivational Factors associated with Parental Involvement Activities

With regard to the association of parental motivation with parental involvement activities, this study found that parents of all groups were highly motivated to invest in education, and also reported to make respective investments. For example, the majority of parents aspired at least a technical diploma or first degree in higher education for their children. That is, regardless of their socioeconomic status, which has been reflected through groups of home resources for learning (HRL), and through language groups in this study. With regard to parental involvement activities, descriptive item statistics showed that parents of all groups were involved with their children's education on a daily or weekly basis. These include activities in both dimensions of parental involvement identified in this study, namely *Supervision* and *Practice*. Hence, the frequency of getting involved is apparently not concerning in either of the considered groups.

Furthermore, analyses of the association between parental motivation with parental involvement activities have been conducted. Results showed a significantly positive effect of parents' perceptions of invitations for involvement from the school on the frequency of home-based parental involvement activities. The finding is true for all groups considered in this study. Hence, regardless of the socioeconomic status, the level of parents feeling included in their children's education by the school is prone to be motivational for home-based involvement activities. Though, the result is somewhat contradicting the general impression of the effect of school invitations for involvement reported in the literature

(e.g., Green et al., 2007; Reiniger & López, 2017; Walker et al., 2011).

With regard to parental educational aspirations, results of this study showed inconsistent findings across pairs. Parental educational aspirations and involvement activities were significantly associated with parental involvement activities of *Practice* for most of the groups and pairs. With regard to the theoretical assumptions of rational decision making, respective findings are rather surprising, given that especially socioeconomic low-performers (i.e., LHRL, Native) were neither expected to put such a strong emphasis on educational attainment, nor to take educational investments. Recall, similar behavioral patterns were also observed in the prefixed qualitative study and resulted in type I *Targeted*, where the value of education itself presumably motivated educational investment behavior. Furthermore, results are in line with the current state of research, finding parental educational aspirations to be the most influential aspect of parental involvement (e.g., Castro et al., 2015; X. Fan & Chen, 2001). However, effect sizes of the observed associations between the motivational factor of parental educational aspirations with parental involvement activities are rather small, and so are shares of explained variance. On that note, results have to be considered with caution.

To conclude, results of the theoretical frameworks' first phase showed that parents' perception of school invitations for involvement were more strongly associated with parental involvement activities, than parental educational aspirations. The overall consideration of parental educational aspirations being the most influential dimension of parental involvement cannot be confirmed for the South African context, and with regard to the considered societal groups. On the contrary, findings rather indicated that parental educational aspirations should not be considered a universal answer to the question on how to motivate parental involvement activities. However, the explanatory power of motivational factors for parental involvement activities found in this study is rather small, especially with regard to the dimension of *Practice*. Hence, further research could work towards the identification of additional factors that potentially capture greater shares of variance in parental involvement activities.

Volitional Aspects associated with Student Outcomes in Learning

With regard to the association of parental involvement activities with student outcomes in learning, this study found that effects of both domains of involvement activities (i.e., *Supervision* and *Practice*) are insignificantly associated with student outcomes in learning for socioeconomic low-performing groups (i.e., LHRL and MHRL). Hence, the positive effect of parental motivation with parental involvement activities, as represented in the first part of the model, is dissolved throughout the process of actional behavior. Although, involvement activities are carried out regularly by socioeconomic low-performing parents, they do not have a significant effect on students' educational performance. In contrast are findings for the group HHRL, where parental involvement activities of *Supervision* are associated significantly positive with student outcomes in learning. Given that this group is considered to be socioeconomically high performing, respective results are in line with the theoretical assumptions of rational decision making; that is, due to risk of loss of status (SV), parents make a greater effort to become involved. Furthermore, they are also prone to having access to resources needed for respective involvement activities to be effective (R. Becker, 2000; Esser, 1999; Stocké, 2010). Results also showed a negative association of parental activities of *Practice* in group HHRL with student outcomes in learning. Hence, parents practicing mathematics or reading with their children foster an overall deterioration of both, reading achievement as well as students' personal motivational beliefs. Similar findings have been obtained for the German and Canadian context with regard to students' overall reading achievement as shown by Feld (2018), who referred to the Self-Determination Theory of Deci & Ryan (2012) in the attempt of explaining respective effects (sec. 3.3). On that note, parental involvement activities of *Practice* may diminish students' autonomy to learn (Feld, 2018).

Although parental involvement activities are generally perceived as a positive and powerful tool to enhance student outcomes in learning (e.g., X. Fan & Chen, 2001; Jeynes, 2005, 2007), and are a constant element of school reform policies (e.g., Domina, 2005; Epstein, 1987), results of this study showed that a more nuanced approach to the exploration of the concept provides a better understanding of its effects. That is, for example,

to consider different domains of parental involvement activities as well as accounting for socioeconomically diverging societal groups in South Africa. In terms of the latter, findings of this study could not confirm the positive effect of home-based parental involvement activities on student outcomes in learning (e.g., Caro, 2011, 2018; Epstein, 1987). For instance, involvement activities of socioeconomic low-performing parents (groups LHRL, MHRL) were not significantly associated with student outcomes in learning. Lastly, and with regard to descriptive statistics reported in the beginning of this section, findings of the multi-group SEM emphasized that the frequency of parental involvement activities is not equal to the quality of parental involvement activities.

A special case presents itself for the pair Native-NonNative, representing language groups in the South African population.

Other than the pair LHRL-MHRL and group LHRL in pair LHRL-HHRL, group Native performed similar to its counterpart (i.e., NonNative), even though group Native is considered to represent lower socioeconomic status. Results showed that parents of group Native were motivated by parents' perceptions of invitations for involvement from the school and parental educational aspirations. Furthermore, parental involvement activities were significantly associated with student outcomes in learning, and effects of parental involvement activities were comparably strong between groups Native and NonNative. That is, actions of *Practice* showed significant negative effects, and actions of *Supervision* showed significant positive effects on student outcomes in learning. Although not directly comparable, similar patterns were observed for group HHRL in the unrestricted model of pair LHRL-HHRL.

Results for pair Native-NonNative are not in line with the assumption that language groups, and specifically native African language groups, function as a proxy for low socioeconomic status. More precisely, results of multi-group SEM rather indicated that the group Native shows educational investment behavior similar to group NonNative. That is, although students of group Native show the lowest average achievement score in reading (294 points) out of all groups considered in this study, parental involvement activities of

Supervision are still significantly positive associated therewith. A possible explanation for this is that students performing at the lowest end of the achievement distribution might be due to the fact that they are educated in a native African language, and therefore prone to be affected by lower quality of instruction, for example through shortage of school resources for learning (Howie et al., 2012). Another reason could be that group Native might also be rather heterogeneous in terms of socioeconomic status, including many parents who actually do not fit the criteria of socioeconomic low-performers, and thus, rather belong to a substantially rising Black-African middle class (e.g., van Staden et al., 2016). The latter may indeed be capable of providing resources for educational investments and effective involvement activities, potentially leading to significant effects in student outcomes in learning as observed for socioeconomic high-performing groups (i.e., HHRL, NonNative).

9.3. Advantages of the Study

Based on the findings presented above, the following prevalent advantages are drawn from this study.

First, the psychological model of action phases, that is the Rubicon Model, as well as the Model of Parental Involvement Process aligned therewith, showed to be a beneficial choice for looking at investment behavior in South Africa. The theoretical framework based on both approaches allowed for the separate consideration of different phases of investment behavior, enabling the identification of action phases at which positive effects dissipated. Particularly striking in this regard is the finding that involvement activities of socioeconomic low-performing parents are not significantly associated with student outcomes in learning, although they are as regularly involved as their counterparts, which are socioeconomic high-performing parents.

Second, whereas the theoretical framework proved to be a beneficial asset for exploring investment behavior within and between societal groups of South Africa, premises of EV-Theory did not capture the reality of educational investment behavior in of the focus

group sufficiently. That was confirmed through findings of both, the qualitative study and the quantitative study, and with regard to turning motivation into action. Evaluated and verified primarily for societal contexts of the U.S. and Europe, the theoretical assumptions of rational decision making leading to educational non-investments among socioeconomic low-performers could not be confirmed for South Africa. That is, results of both studies, and hence, with regard to overall as well as parental educational investment behavior among socioeconomic low-performers, indicating high levels of motivation as well as volitional behavior with regard to educational investments.

Third, particular advantages result from the quantitative database utilized for this study. That is, prePIRLS provides a useful data set when aiming to validate theoretical concepts across education systems. This is due to the standardization of the assessment, and its comparability (van Staden & Howie, 2014). However, this study showed that respective comparisons are not only possible between but also within education systems. Providing a representative oversample of n = 15,744 students to map language patterns in South Africa, prePIRLS 2011 enabled analyses of large and presumably homogeneous groups of the South African population, representing the lower and upper end of the socioeconomic distribution, while allowing to neglect the middle class. Along with this, the sample size of groups also enabled the application of advanced multivariate statistical analyses, that are multi-group SEMs. To enhance the quality of the models, this study furthermore summarized large numbers of missing cases on the scale of home resources for learning (HRL) to a separate group, namely cases with missing information (MHRL). This was done to obtain information of this group's educational investment behavior compared to the reference group (LHRL). To that end, comparability of measurement models, that is (partial) measurement invariance on domains of parental involvement activities (*Supervision*, *Practice*), was ensured for all pairs utilized in this study. Results revealed similar patterns of volitional educational behavior of parents who omitted information on the socioeconomic status, and those who were identified to have low socioeconomic status. Hence, the utilized procedure provided valuable information on behavioral patterns of those who would otherwise be excluded from the analyses, or ascribed with a socioe-

conomic status through multiple imputation of missing information based on auxiliary variables (Lüdtke et al., 2007).
Lastly, the application of a qualitative study followed by a quantitative study allowed for the exploration of educational investment behavior from two different perspectives and with regard to two different samples of socioeconomic low-performers in South Africa. These are, first, adults investing in their own education represented by a non-representative sample of street vendors in Cape Town/South Africa, and second, a representative sample of parents investing in their children's education. Results of the qualitative study are considered particularly noteworthy, given that they disclosed reasons for non-investments as well (i.e., the dilemma earning vs. learning).

9.4. Limitations of the Study

Alongside the aforementioned advantages, this study also showed some noteworthy limitations.

First, limitations result from the study design incorporating multiple methods rather than mixing them. As argued in section 6, a sequence of qualitative and quantitative studies alone does not make a mixed-methods design (Creswell & Plano Clark, 2007). Although both studies incorporated here are considered to be complementary regarding the question of educational investment behavior among socioeconomic low-performers in South Africa, the mixing thereof is limited due to different samples and objectives of each study. Arguably, a true mixed-method design expanding the quantitative model by a qualitative perspective could provide more targeted information. That is, for example, with regard to the question on the quality of home-based involvement activities performed by socioeconomic low-performing parents.

With regard to the qualitative study, limitations mainly concern the richness of information on the topic of educational investment behavior obtained from the qualitative interview, as well as the study administration. That is, in addition to a rather small

sample size, problem-centered interviews had to be conducted in English rather than the street vendors' mother tongue (sec. 7.1). This was due to limited resources for administering the study, and to ensure adequate communication between the interviewer and the interviewees. As a result thereof, interview transcripts offered limited informational content. Hence, conducting the interviews in the respective mother tongue of the interviewee, and have them translated into English afterwards could potentially enhance the richness of informational content, and consequently allow for more substantiated data analysis.

For the quantitative perspective, the most obvious limitation results from secondary analysis on cross-sectional data from large-scale assessments such as prePIRLS.
First, secondary analysis on preexisting data is always limited with regard to the availability of information. For this study, this means that measures of home-based parental involvement are merely an approximation of those originally intended for the operationalization of the dimension as appointed in the revised Model of Parental Involvement Process (Hoover–Dempsey & Sandler, 1997; Hoover–Dempsey et al., 1992, 2005; Walker et al., 2005). Furthermore, information utilized for the operationalization of home-based parental involvement were reported by the parents themselves, with the home background questionnaire of prePIRLS. Peng & Wright (1994) pointed to issues of biased information that might occur from self-reported data with regard to parental involvement activities. Non-experimental information as drawn from prePIRLS can therefore only serve as an approximate impression of effects (Peng & Wright, 1994).
Second, cross-sectional data does not allow for statements on the causality of effects. For example, whether parents' perceptions of invitations for involvement from the school continuously fosters home-based parental involvement activities among socioeconomic low-performing parents in South Africa cannot be answered from this study. Rather, this study provides snapshots of the association between motivational and volitional factors of educational investment behavior, and thus, provides meaningful points of reference for further analyses. For example, with regard to trends in the effects of parental involvement activities on country level (sec. 10).

Third, this study incorporated scales derived from manifest items on reading achievement, students' intrinsic motivation as well as self-efficacy in reading. The main reason for that was to ensure comparability between studies, and to ensure the usage of a quality indicator built according to the standards of psychometric scaling methodologies (sec. 8.2.3). However, with regard to the context of South Africa, scaling of respective constructs within the SEM could account for country-specific adjustments.[69]

In addition to limitations resulting from the data set, some constraints also result from design decisions taken for the quantitative study.
For instance, this study accounted for different societal groups. To that end, different items were utilized, aimed at reflecting on socioeconomic status. Groups consisted of those with access to few and many home resources for learning (LHRL, HHRL), as well as those who omitted respective information (MHRL). Furthermore, language groups (Native, NonNative) were derived from a combination of items. Though, with regard to the latter, results of this study pointed to the need of a more nuanced solution for grouping. That is, although the segregation of the South African population is continuously described through discrepancies in languages, and more precisely in native versus non-native language groups (e.g., Howie et al., 2012; van Staden & Howie, 2012, 2014), recent research also argues for more distinct categories, differentiating between English and Afrikaans, as well as grammatically and linguistically homogeneous groups of native African languages (e.g., van Staden, 2010; van Staden & Howie, 2012; van Staden et al., 2016). However, with regard to home resources for learning in association with the language of test, findings on prePIRLS 2011 showed that the majority of students exposed to many home resources for learning was assessed in English or Afrikaans. On the contrary, only a few students assessed in isiNdebele, siSwati, Tshivenda and Xitsonga had access to alike (van Staden et al., 2016). Further analyses would need to explore whether more

[69] For example, non-referenced pre-analyses of confirmatory factor analyses suggested that parents' educational and occupational status included to the scale of home resources for learning (HRL) are highly correlated in South Africa, and hence, issues of local item dependence would need to be explored and potentially addressed.

distinct language groups lead to a clearer picture of societal groups and their educational investment behavior.

Finally, limitations occur from the multi-group SEM itself. The model was built according to the premises of the theoretical framework utilized for this study. In order to keep the model parsimonious, only information considered theoretically relevant for the exploration of parental educational investment behavior were included. However, this methodological decision was accompanied by leaving potentially relevant control variables such as students' age untended (e.g., van Staden & Howie, 2012).

Lastly, the explanatory power of the model needs to be addressed. Overall, the model explained around 10% of variance in parental involvement. Differences in the average frequency of parental activities of *Supervision* are explained slightly better through motivational factors, than parental activities of *Practice*. Furthermore, parental involvement activities of *Supervision* and *Practice* explained rather small shares of variance in reading achievement best for socioeconomic high-performing groups (i.e., 17.4% in HHRL and 11.7% in Native). Overall, the explanatory power of the model is rather limited, and hence, results need to be considered with caution.

10. Summary & Outlook

Finally, this section briefly summarizes the major findings of this study, and points towards potential follow-up questions for research.

10.1. Summary of Results

Introducing this study, photographic work by Johnnie Miller was used as an example to visualize the persisting segregation and socioeconomic inequality in South Africa more than two decades after the Apartheid regime officially came to an end. Along those lines, the question was posed as to which factors can counteract respective societal patterns. Arguing from the perspective of Max Weber (1922), who identified educational attainment as the most decisive factor for the enhancement of upward mobility and life opportunities, this study took a closer look at patterns of educational investment behavior in consideration of different societal groups of the South African population. At the core of this study was the exploration of investment behavior among socioeconomic low-performers. To analyze thereof, a qualitative and a quantitative perspective was utilized. Referring back to the research questions I - III of this study (sec. 5), discussed results (sec. 8.3) are concluded in the following.

From a general perspective, it can be summarized that socioeconomic low-performers in South Africa are motivated to invest in education (I), and that they selectively turn respective motivation into actions (II). That is true for street vendors acting according to type I and investing in their own education, as well as for the group of socioeconomic low-performing parents being actively involved in their children's education. However, respective findings contradict the premises of EV-Theory. Lastly, educational investments of South African socioeconomic low-performing parents do not have a significant effect on students' outcomes in learning (III). That is, except for group Native. Possible explanations as well as solutions have been discussed in section 9.

Finally, it was concluded that the theoretical assumptions for motivational and actional behavior of socioeconomic low-performers based on EV-Theory were not explicitly reflected by empirical results of this study; and that socioeconomic low-performers in South

Africa lack resources such as time and money as well as resources for highly effective involvement rather than the motivation for educational investments itself.

10.2. Implications of the Study

Subsequent to the main findings of this study, its implications are summarized. In this respect, two elevating findings with regard to understanding and supporting educational investment behavior among socioeconomic low-performers are noteworthy.

First, findings showed that socioeconomic low-performers in South Africa are highly motivated to invest in education, that is they acknowledge the importance of educational attainment for the improvement of life opportunities, as well as upward mobility. However, non-investments are often due to the lack of time to invest in education. That has been reflected by the dilemma earning vs. learning within the qualitative perspective of this study. Considering the discussed limitations of this study, as well as the need for further research, South African policies could focus on utilizing the motivational state of socioeconomic low-performers, and aim at facilitating educational investments, for example, through the basic provision of social security that enables educational attainment without risking loss of income. Respective strategies might also be in line with the request for an adaptation of the *Broad-Based Black Economic Empowerment* Act (BBBEE) (sec. 2.2.3), for example, in terms of a "broader definition of empowerment moving away from the limited focus of ownership" (Shai et al., 2019, p. 20). Though, the BBBEE considerably added to the transformation of societal patterns already (Patel & Graham, 2012), incorporating the realization of educational investments might be a meaningful addition, given that thereof could potentially result in highly qualified personnel. The therewith considered causality has been acknowledged by researchers already, stating that "economic empowerment has to be premised on a transformation of education and skills capacity building" (Horwitz, 2011, p. 314). Results of this study support respective voices.

Second, along the lines of empowerment policies are also implications with regard to parental involvement as a form of educational investment. In that respect, policy makers could acknowledge that the goal must not be to primarily enhance the motivation for

parental educational investments or its frequency, but to foster its quality. For instance, governments could, first, enable an empirical investigation of the quality of home-based parental involvement, especially with regard to activities of *Supervision*, as those have shown to be more beneficial to student outcomes in learning; and second, attempt to establish support programs for parents to help accommodate effective situations of home-based parental involvement. Building the bridge to the aforementioned, effective parental involvement may be positively associated with student outcomes in learning, which has been confirmed already within this study for the group of socioeconomic high-performers and with regard to activities of *Supervision*. Surely, well supported, well performing, and hence, academically well-prepared students are an asset to the success of programs such as the aforementioned BBBEE (sec. 2.2.3). That is, the better students are equipped with skills and knowledge obtained from high-quality education, the better they may be equipped for managerial positions or ownership. Moreover, governmental efforts could continue to aim for equal learning opportunities for all, and thus, support the mitigation of socioeconomically caused differences between student outcomes in learning. For instance, the attempt of no-fee schools, and the introduction of compulsory pre-primary education (sec. 2.2.4) are positive examples for such attempts. Respective thoughts are also reflected in the opening statement of this study: "Because there is nothing we can do without education" (Street vendor in Cape Town/South Africa, 2013).
However, respective policy recommendations must be considered with caution, given that the explanatory power of this study is compromised due to several limitations (sec. 9.4).

10.3. Recommendations for further Research

Lastly, the following suggestions for further research are tied to the results and implications presented above.

This study solemnly focused on student outcomes in learning with regard to reading and literacy. Aimed at substantiating the results of this study, students' performance in additional subject domains such as mathematics and science, as well as their personal

motivational beliefs could be considered. A suitable study for this purpose is TIMSS (sec. 2.2.4). South Africa has participated in the study since 1995 with grade 5 and grade 8, although data for fifth-grade students is only available for the cycles of 2015 and 2019. It is recommended to prioritize younger students, given that parental involvement activities are likely to be more effective at an earlier age (e.g., Caro, 2011).

Utilizing data from prePIRLS 2011 and thereby ensuring a timely link between the qualitative and quantitative intake, this study necessarily disregarded developments of the country over the last decade. To substantiate the findings of this study and account for current developments in South Africa further, analyses of this study should be replicated with more recent data. Suitable options are TIMSS 2015 and 2019, PIRLS and PIRLS Literacy 2016, and latest data of PIRLS 2021. Similar to prePIRLS 2011, South African data of PIRLS Literacy 2016 consists of a student oversample as well (Howie et al., 2017), allowing for the analysis of population groups large enough in size to ensure the reliable performance of advanced statistical models such as SEMs.

Furthermore, this study utilized cross-sectional data, disabling statements on the causality of effects. Hence, trend analyses at country level, focusing the development of student outcomes in learning in association with parental involvement activities between different cycles of TIMSS or PIRLS would add informative insights in this regard.[70]

With regard to the study design, two recommendations for further research are derived. For example, concerning the theoretical framework of this study as well as empirical findings related therewith, a potential follow-up question would be whether respective concepts of EV-Theory utilized to substantiate the framework of human action behavior could be adapted in a way to fulfill the purpose of a generalizable theoretical framework.

[70] It must be noted that reported trends in South African students' overall reading achievement between the cycles of prePIRLS 2011 and PIRLS Literacy 2016 cannot be considered reliable due to errors in recalibration of the data (Gustafsson, 2020). Instead, micro trend data for both cycles should be considered, and viewed in association with potential factors underlying the nature of growth. To that end, it is suggested to consider socioeconomic status as well as language groups (van Staden, 2020).

That is, capturing international contexts exceeding the theory's primary scope of the U.S. and Europe. For this purpose, additional research exploring the validity of respective theoretical assumptions in different educational systems is needed. Particular attention could be paid to countries that show patterns of social segregation as well, for example Mexico and Brazil. Possibly, the findings of this study could be reinforced, implying a reconsideration of the general assumption that cost-benefit calculations of socioeconomic low-performers are likely to lead to educational non-investments.

Furthermore, this study took a qualitative and a quantitative perspective on the subject of educational investment behavior. The qualitative perspective of this study is considered particularly valuable, given that it disclosed reasons for non-investments. However, similar explanatory results, for example on the question how the quality of parental involvement activities differs between socioeconomic low-performers and high-performers, could not be gathered from the standardized quantitative data set of prePIRLS 2011. To that end, the strategic inclusion of semi-open question formats on particular country-specific subjects within the national versions of background questionnaire of large-scale assessments could potentially enhance its informative value and should be explored accordingly.

Lastly, by focusing educational investment behavior and parental involvement activities as a form thereof, this study addressed an important aspect of the education system (e.g., Domina, 2005; X. Fan & Chen, 2001). Regardless of the effect that researchers and policy makers ascribed parental involvement with in the past, it must be acknowledged that thereof is considered only one aspect of an effective learning environment, which may or may not shape educational careers of students, and enhance upward mobility in the long run. That is, although parental involvement needs to be considered "an important ingredient for the remedy for many problems in education" (X. Fan & Chen, 2001, p. 1), it is only that: one ingredient among many.

Hence, the key finding of this study, namely the need for exploring and supporting quality of parental involvement activities among socioeconomic low-performers to potentially enhance student outcomes in learning, is only a piece of the puzzle in the attempt of

counteracting social inequality as caused by historical and political events in the case of South Africa. Especially with regard to the remaining socioeconomic segregation of the South African population (sec. 2), a central task for further research is the expansion of this study's scope towards the inclusion of additional impact factors, for example at teacher-, school-, and curriculum level. For the purpose of identifying thereof, further research could certainly utilize well-acknowledged generic frameworks, reflecting on the multiplicity of impact factors influencing educational outcomes. The consideration of a country's cultural context is crucial for the identification of such factors (Bos, 1999). With regard to the specific case of South Africa, connecting points potentially emerge from findings confirming a highly segregated society, where Black-African students are still subjected to a less optimal student-teacher ratios, and are educated in schools with less resources for learning (e.g., Howie et al., 2012, 2017). Hence, country-specific recommendations could be to control for different societal groups, as performed in this study, and to consider school-level factors potentially reinforces the multiplication of inequalities in educational opportunities. To that end, multilevel multi-group structural equation modeling is recommended.

References

Alderfer, C. P. (1969). An empirical test of a new theory of human needs. *Organizational Behavior and Human Decision Processes*, *4*(2), 142–175.

Anderson, K. J. (2005). *Understanding parents' decision to become involved in their childrens' education* [Doctoral dissertation, University of Delaware]. Retrieved August 30, 2022, from https://www.proquest.com

Ardington, C., Branson, N., Lam, D., & Leibbrandt, M. (2011). Explaining the persistence of racial gaps in schooling in South Africa. *African Population Studies*, *25*(2), 509–542.

Asparouhov, T. (2017). *Difference testing using chi-square* [Mplus Discussion/Structural Equation Modeling]. Retrieved August 30, 2022, from http://www.statmodel.com/discussion/messages/11/8230.html?1590018965

Asparouhov, T., & Muthén, B. O. (2022). *Multiple imputation with Mplus* [Version 4]. Retrieved August 30, 2022, from https://www.statmodel.com/download/Imputations7.pdf

Atkinson, J. W. (1964). *An introduction to motivation*. Van Nostrand.

Attewell, P., & Newman, K. S. (2010). *Growing gaps. Educational inequality around the world*. Oxford University Press.

Bakker, J., Denessen, E., & Brus–Laeven, M. (2007). Socio-economic background, parental involvement and teacher perceptions of these in relation to pupil achievement. *Educational Studies*, *33*(2), 177–192.

Bandalos, D. L. (2014). Relative performance of categorical diagonally weighted least squares and robust maximum likelihood estimation. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(1), 102–116.

Barnard, W. M. (2004). Parental involvement in elementary school and educational attainment. *Children and Youth Services Review*, *26*(1), 39–62.

Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling: A Multidisciplinary Journal*, *13*(2), 186–203.

Becker, G. S. (1962). Investment in human capital. A theoretical analysis. *The Journal of Political Economy*, *70*(5.2), 9–49.

Becker, R. (2000). Klassenlage und Bildungsentscheidungen. Eine empirische Anwendung der Wert-Erwartungstheorie. [Class position and educational choices. An empirical application of Value-Expectancy Theory.] *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, *52*(3), 450–474.

Becker, R. (2003). Educational expansion and persistent inequalities of education. Utilizing subjective expected utility theory to explain increasing participation rates in upper secondary school in the federal republic of Germany. *European Sociological Review*, *19*(1), 1–24.

Becker, R., & Lauterbach, W. (2016). Bildung als Privileg. Ursachen, Mechanismen, Prozesse und Wirkungen. [Education as privilege. Causes, mechanisms, processes and effects.] In R. Becker & W. Lauterbach (Eds.), *Bildung als Privileg. Erklärungen und Befunde zu den Ursachen der Bildungsungleichheit. [Education as privilege. Explanations and findings on the causes of educational Inequality.]* (pp. 3–53). Springer VS.

Bempechat, J., & Shernoff, D. J. (2012). Parental influences on achievement motivation and student engagement. In S. Christenson, A. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 315–342). Springer.

Bentler, P. M., & Chou, C.-P. (1987). Practical issues in structural modeling. *Sociological Methods & Research*, *16*(1), 78–117.

Bernstein, B. B. (1975). *Class, codes and control. Towards a theory of educational transmissions*. Routledge.

Bernstein, H. (1972). Schools for servitude. In A. La Guma (Ed.), *Apartheid. A collection of writings on South African racism by South Africans* (pp. 43–80). Lawrence and Wishart.

Bloom, B. S. (1980). The new direction for educational research: Alterable variables. *Phi Delta Kappan*, *61*(6), 382–385.

Bojuwoye, O., & Narain, M. (2008). Parental involvement and children's academic achievement in a South Africa setting. *Journal of Psychology in Africa*, *18*(2), 275–278.

Bong, M. (2008). Effects of parent-child relationships and classroom goal structures on motivation, help-seeking avoidance, and cheating. *The Journal of Experimental Education*, *76*(2), 191–217.

Bos, W. (1999). Empirisch-vergleichende Forschung. Konzeptionelle Probleme und politischer Nutzen. [Empirical-comparative research. Conceptual problems and policy benefits.] *Tertium Comparationis*, *5*(2), 118–133.

Boudon, R. (1974). *Education, opportunity, and social inequality. Changing prospects in western society*. Wiley.

Bourdieu, P., & Passeron, J.-C. (1977). *Reproduction in education, society, and culture*. Sage.

Branson, N., Garlick, J., Lam, D., & Leibbrandt, M. (2012). *Education and inequality: The South African case* (SALDRU, Ed.) [A Southern Africa Labour and Development Research Unit Working Paper Number 75]. University of Cape Town. Retrieved August 30, 2022, from https://www.opensaldru.uct.ac.za/bitstream/handle/11090/168/2012_75.pdf?sequence=1

Branson, N., Hofmeyer, C., & Lam, D. (2012). *The impact of the no-fee policy on enrolement and school performance. Evidence from NIDS waves 1-3* (SALDRU, Ed.) [A Southern Africa Labour and Development Research Unit Working Paper Number 197]. University of Cape Town. Retrieved August 30, 2022, from https://www.opensaldru.uct.ac.za/bitstream/handle/11090/856/2017_197_Saldruwp.pdf?sequence=1

Breen, R., & Goldthorpe, J. H. (1997). Explaining educational differentials: Towards a formal rational action theory. *Rationality and Society*, *9*(3), 275–305.

Britannica. (n.d.). *Bantustan* (Encyclopedia Britannica, Ed.). Retrieved August 30, 2022, from https://www.britannica.com/topic/Bantustan

Buchmann, C. (2002). Measuring family background in international studies of education. Conceptual issues and methodological challanges. In A. C. Porter & A. Gamoran

(Eds.), *Methodological advances in cross-national surveys of educational achievement* (pp. 150–197). National Acadamy Press.

Budlender, D. (2011). *Statistics on informal employment in South Africa* (WIEGO, Ed.) [WIEGO Statistical Brief No. 3]. Retrieved August 30, 2022, from https://www.wiego.org/sites/default/files/publications/files/Budlender_WIEGO_SB3.pdf

Bunting, B. (1972). The origings of Apartheid. In A. La Guma (Ed.), *Apartheid. A collection of writings on South African racism by South Africans* (pp. 23–40). Lawrence and Wishart.

Burger, R., van der Berg, S., van der Walt, S., & Yu, D. (2015). *The long walk. Considering the enduring spatial and racial dimension of deprivation two decades after the fall of apartheid* (University of Stellenbosch, Department of Economics, Ed.) [Stellenbosch Economic Working Papers 17/15]. Retrieved August 30, 2022, from https://www.ekon.sun.ac.za/wpapers/2015/wp172015

Byrne, B. M. (2012). *Structural equation modeling with Mplus. Basic concepts, applications, and programming* [Multivariate Applications Series]. Routledge.

Byrne, B. M., Shavelson, R. J., & Muthén, B. O. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, *105*(3), 456–466.

Campo, J. (2011). *Parental involvement in education: Model exploration among parents of elementary and middle school students* [Doctoral dissertation, Saint Louis University]. Retrieved August 30, 2022, from https://www.proquest.com

Caro, D. H. (2011). Parent-child communication and academic performance. Associations at the within- and between-country level. *Journal for Educational Research Online*, *3*(2), 15–37.

Caro, D. H. (2018). Socio-economic gaps in subject interest: the mediating role of parental cognitive involvement. *Large-Scale Assessments in Education*, *6*(13), 1–38.

Caro, D. H., Sandoval-Hernandez, A., & Lüdtke, O. (2014). Cultural, social, and economic capital constructs in international assessments: an evaluation using exploratory

structural equation modeling. *School Effectiveness and School Improvement*, *25*(3), 433–450.

Castro, M., Expósito-Casas, E., López-Martín, E., Lizasoain, L., Navarro, E., & Gaviria, J. L. (2015). Parental involvement on student academic achievement. A meta-analysis. *Educational Research Review*, *14*, 33–46.

Catsiapis, G. (1987). A Model of educational investment decisions. *The Review of Economics and Statistics*, *69*(1), 33–41.

Chen, F., Sakyi, A., & Cui, Y. (2021). Linking student, home, and school factors to reading achievement. The mediating role of reading self-efficacy. *Educational Psychology*, *41*(10), 1260–1279.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464–504.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(2), 233–255.

Cheung, W. M., Lam, J. W. I., Au, D. W. H., So, W. W. Y., Huang, Y., & Tsang, H. W. H. (2017). Explaining student and home variance of Chinese reading achievement of the PIRLS 2011 Hong Kong. *Psychology in the Schools*, *54*(9), 889–904.

Chiu, M. M. (2010). Effects of inequality, family and school on Mathematics achievment. Country and student differences. *Social Forces*, *88*(4), 1645–1676.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Jorunal of Sociology*, *94*, 95–120.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*. US Congressional Printing Office.

Combrinck, C., van Staden, S., & Roux, K. (2014). Developing early readers. Patterns in introducing critical reading skills and strategies to South African children. *Reading & Writing*, *5*(1), 1–9.

Conley, D. (2001). Capital for college. Parental assets and postsecondary schooling. *Sociology of Education*, *74*(1), 59–72.

Creswell, J. W. (2016). *Educational research: Planning, conducting, and evaluating quantiative and qualitative research*. Pearson.

Creswell, J. W., & Plano Clark, V. L. (2007). *Designing and conducting mixed methods research*. Sage.

Desforges, C., & Abouchaar, A. (2003). *The impact of parental involvement, parental support and family education on pupil achievements and adjustment. A literature review* [Research Report RR433]. Department for Education and Skills. Retrieved August 30, 2022, from https://dera.ioe.ac.uk/6305/1/rr433.pdf

Desland, R., & Betrand, R. (2005). Motivation of parent involvememt in secondary-level schooling. *The Journal of Educational Research*, *98*(3), 164–175.

Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development*, *43*(1), 121–149.

Domina, T. (2005). Leveling the home advantage. Assessing the effectiveness of parental involvement in elementary school. *Sociology of Education*, *78*(3), 233–249.

Dumont, H., Trautwein, U., Lüdtke, O., Neumann, M., Niggli, A., & Schnyder, I. (2012). Does parental homework involvement mediate the relationship between family background and educational outcomes? *Contemporary Educational Psychology*, *37*(1), 55–69.

Eccles, J. S., Wigfield, A., & Schiefele, U. (1998). Motivation to succeed. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (pp. 1017–1095). Wiley.

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *8*(3), 430–457.

Epstein, J. L. (1986). Parents' reactions to teacher practices of parent involvement. *The Elementary School Journal*, *86*(3), 277–294.

Epstein, J. L. (1987). Parent involvement. What research says to administrators. *Education and Urban Society*, *19*(2), 119–136.

Epstein, J. L. (1992). School and family partnerships. In M. Aikan (Ed.), *Encyclopedia of educational research* (pp. 1139–1151). MacMillan.

Epstein, J. L. (1995). School/family/community patnerships. Caring for the children we share. *Phi Delta Kappan*, *76*(9), 701–712.

Epstein, J. L., & Salinas, K. C. (1992). *Promising programs in the middle grades*. National Association of Secondary School Principals.

Erikson, R., & Jonsson, J. O. (1996). Explaining class inequality in education. The Swedish test case. In R. Erikson & J. O. Jonsson (Eds.), *Can education be equalized? The Swedish case in comparative perspective* (pp. 1–63). Westview Press.

Esser, H. (1999). *Soziologie. Spezielle Grundlagen. [Sociology. Particular foundations.] 1: Situationslogik und Handeln*. Campus.

Fan, W., & Williams, C. M. (2010). The effects of parental involvement on students' academic self-efficacy, engagement and intrinsic motivation. *Educational Psychology*, *30*(1), 53–74.

Fan, X., & Chen, M. (2001). Parental involvement and students' academic achievement. A meta-analysis. *Educational Psychology Review*, *13*(1), 1–22.

Fan, X., & Sivo, S. A. (2005). Sensitivity of fit-indexes to misspecified structural or measurement model components. Rationale of two-index strategy revisted. *Structural Equation Modeling: A Multidisciplinary Journal*, *12*(3), 343–367.

Feld, I. (2018). *Parental involvement and social background in Canada and Germany. Limitations and possibilities of a comparative analysis of the Progress in International Reading and Literacy Study (PIRLS) 2011* [Doctoral dissertation]. Waxmann.

Felix, N., Dornbrack, J., & Scheckle, E. (2008). Parents, homework and socio-economic class. Discourses of deficit and disadvantage in the "new" South Africa. *English Teaching: Practice and Critique*, *7*(2), 99–112.

Fetters, M. D., & Molina-Azorin, J. F. (2017). The journal of mixed methods research starts a new decade. Principles for bringing in the new and divesting of the old language of the field. *Journal of Mixed Methods Research*, *11*(1), 3–10.

Flick, U. (2012). *Qualitative Sozialforschung. Eine Einführung [Qualitative research. An introduction.]* Rowohlt Taschenbuch Verlag.

Foy, P., Brossman, B., & Galia, J. (2012). *Methods and procedures in TIMSS and PIRLS 2011: Scaling the TIMSS and PIRLS 2011 Achievement Data* (International Study Center Lynch School of Education, Boston College, Ed.). IEA. Retrieved August 30, 2022, from https://timssandpirls.bc.edu/methods/pdf/tp11_scaling_achievement.pdf

Foy, P., & Drucker, K. T. (2013a). *International version of the PIRLS 2011 background questionnaires and curriculum questionnaire* (International Study Center Lynch School of Education, Boston College, Ed.) [Supplement 1]. IEA. Retrieved August 30, 2022, from https://timssandpirls.bc.edu/pirls2011/downloads/P11_UG_Supplement1.pdf

Foy, P., & Drucker, K. T. (2013b). *PIRLS 2011 User Guide for the international database* (International Study Center Lynch School of Education, Boston College, Ed.). IEA. Retrieved August 30, 2022, from https://timssandpirls.bc.edu/pirls2011/downloads/P11_UserGuide.pdf

Geiser, C. (2011). *Datenanalyse mit Mplus. Eine anwendungsorientierte Einführung. [Data analysis with Mplus. An application-oriented introduction.]* Springer VS.

Gollwitzer, P. M. (1990). Action phases and mind-sets. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition. Foundations of social behavior* (pp. 53–92). Guilford.

Gonzalez, E., & Kennedy, A. M. (2003). *PIRLS 2001 User Guide for the international database* (International Study Center Lynch School of Education, Boston College, Ed.). IEA. Retrieved August 30, 2022, from https://timssandpirls.bc.edu/pirls2001i/pdf/UserGuide.pdf

Graham, J. W. (2003). Adding missing-data-relevant variables to FIML based structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *10*(1), 80–100.

Green, C. L., Walker, C., Hoover–Dempsey, K. V., & Sandler, H. M. (2007). Parents' motivations for involvement in children's education. An empirical test of a theoretical model of parental involvement. *Journal of Educational Psychology*, *99*(3), 532–544.

Griffith, J. (1998). The relation of school structure and social environment to parent involvement in elementary school. *The Elementary School Journal*, *99*(1), 53–80.

Grolnick, W. S., & Slowiaczek, M. L. (1994). Parents' involvement in children's schooling. A multidimensional conceptualization and motivational model. *Child Development*, *65*(1), 237–252.

Gustafsson, M. (2020). *A revised PIRLS 2011 to 2016 trend for South Africa and the importance of analysing the underlying microdata* (University of Stellenbosch, Department of Economics, Ed.) [Stellenbosch Economic Working Papers WP02/2020]. Retrieved August 30, 2022, from https://www.ekon.sun.ac.za/wpapers/2020/wp022020

Haferburg, C., & Osmanovich, A. (2017). Sozio-öknomische Komplexität Südafrikas. [Socio-economic complexity of South Africa.] In D. de La Fontaine, F. Müller, C. Hofmann, & B. Leubolt (Eds.), *Das politische System Südafrikas. [The political system of South Africa.]* (pp. 53–71). Springer VS.

Hayes, A. F., & Coutts, J. J. (2020). Use Omega rather than Cronbach's Alpha for estimating reliability. But. *Communication Methods and Measures*, *14*(1), 1–24.

Heaton, J. (2008). Secondary analysis of qualitative data. An overview. *Historical Social Research*, *33*(3), 33–45.

Heckhausen, H., & Gollwitzer, P. M. (1987). Thought contents and cognitive functioning in motivational versus volitional states of mind. *Motivation and Emotion*, *11*(2), 101–120.

Heckhausen, H., & Kuhl, J. (1985). From wishes to action. The dead ends and short cuts on the long way to action. In M. Frese & J. Sabini (Eds.), *Goal-directed behavior. Psychological theory and research on action* (pp. 134–160). Erlbaum.

Heckhausen, J., & Heckhausen, H. (2008). *Motivation and action*. Springer.

Heintz, J., & Posel, D. (2008). Revisiting informal employment and segmentation in the South African labour market. *South African Journal of Economics*, *76*(1), 26–44.

Hill, N. E., & Tyson, D. F. (2009). Parental involvement in middle school. A meta-analytic assessment of the strategies that promote achievement. *Developmental Psychology*, *45*(3), 740–763.

Hoover–Dempsey, K. V., Bassler, O. C., & Brissie, J. S. (1992). Explorations in parent-school relations. *The Journal of Educational Research*, *85*(5), 287–294.

Hoover–Dempsey, K. V., Battiato, A., Walker, J. M. T., Reed, R. P., DeJong, J. M., & Jones, K. P. (2001). Parental involvement in homework. *Educational Psychologist*, *36*(3), 195–209.

Hoover–Dempsey, K. V., & Sandler, H. M. (1997). Why do parents become involved in their children's education? *Review of Educational Research*, *67*(1), 3–42.

Hoover–Dempsey, K. V., Walker, J. M. T., Sandler, H. M., Whetsel, D., Green, C. L., Wilkins, A. S., & Closson, K. (2005). Why do parents become involved? Research findings and implications. *The Elementary School Journal*, *106*(2), 105–130.

Horwitz, F. M. (2011). An assessment of employment equity and Broad Based Black Economic Empowerment developments in South Africa. *Equality Diversity and Inclusion: An International Journal*, *30*(4), 297–317.

Houserman, M. (2016). The educational system of South Africa. An overview. *IEM Spotlight Newsletter*, *13*(3).

Howie, S. J. (2003). Language and other background factors affecting secondary pupils' performance in Mathematics in South Africa. *African Journal of Research in Mathematics, Science and Technology Education*, *7*(1), 1–20.

Howie, S. J., Combrinck, C., Roux, K., Tshele, M., Mokoena, G. M., & McLeod Palane, N. (2017). *PIRLS Literacy 2016. South African highlights report* (University of

Pretoria, Ed.). Center for Evaluation and Assessment. Retrieved August 30, 2022, from https://repository.up.ac.za/bitstream/handle/2263/66185/Combrinck_Pirls_2017.pdf?sequence=1&isAllowed=y

Howie, S. J., Combrinck, C., Tshele, M., Roux, K., McLeod Palane, N., & Mokoena, G. M. (2018). *PIRLS 2016 Progress in International Reading and Literacy Study 2016 grade 5 benchmark participation. South African children's reading literacy achievement* (University of Pretoria, Ed.). Center for Evaluation and Assessment. Retrieved August 30, 2022, from https://www.up.ac.za/media/shared/164/ZP_Files/grade_5_pirls_2016_final_country_report.zp147996.pdf

Howie, S. J., & Scherman, V. (2008). The achievement gap between science classrooms and historic inequalities. *Studies in Educational Evaluation*, *34*(2), 118–130.

Howie, S. J., van Staden, S., Tshele, M., Dowse, C., & Zimmermann, L. (2012). *PIRLS 2011. South African children's reading literacy achievement. Summary report* (University of Pretoria, Ed.). Center for Evaluation and Assessment. Retrieved August 30, 2022, from https://repository.up.ac.za/handle/2263/65996

Howie, S. J., Venter, E., & van Staden, S. (2008). The effect of multilingual policies on performance and progression in reading literacy in South African primary schools. *Educational Research and Evaluation*, *14*(6), 551–560.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis. Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55.

IEA's Progress in Reading Literacy Study. (n.d.). *PIRLS 2011 Copyright (c) 2012 International Association for the Evaluation of Educational Achivement (IEA)* (TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College, Ed.). Retrieved August 30, 2022, from https://timssandpirls.bc.edu/pirls2011/index.html

Jeynes, W. H. (2003). A meta-analysis. The effects of parental involvement on minority children's academic achievement. *Education and Urban Society*, *35*(2), 202–218.

Jeynes, W. H. (2005). A meta-analysis of the relation of parental involvement to urban elementary school student academic achievement. *Urban Education*, *40*(3), 237–269.

Jeynes, W. H. (2007). The relationship between parental involvement and urban secondary school student academic achievement. A meta-analysis. *Urban Education*, *42*(1), 82–110.

Jodl, K. M., Michael, A., Malanchuk, O., Eccles, J. S., & Sameroff, A. (2001). Parents' roles in shaping early adolescents' occupational aspirations. *Child Development*, *72*(4), 1247–1265.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research. A research paradigm whose time has come. *Educational Researcher*, *33*(7), 14–26.

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, *1*(2), 112–133.

Joncas, M. (2012). *Methods and procedures in TIMSS and PIRLS 2011: PIRLS 2011 population coverage and exclusions* (International Study Center Lynch School of Education, Boston College, Ed.). IEA. Retrieved August 30, 2022, from https://timssandpirls.bc.edu/methods/pdf/P11_Pop_Coverage.pdf

Joncas, M., & Foy, P. (2012). *Methods and procedures in TIMSS and PIRLS 2011: Sample design in TIMSS and PIRLS* (International Study Center Lynch School of Education, Boston College, Ed.). IEA. Retrieved August 30, 2022, from https://timssandpirls.bc.edu/methods/pdf/tp_sampling_design.pdf

Junge, B. (2017). Bildungspolitik in Südafrika als Spiegelbild von Apartheidideologie und Transformation. [Education policy in South Africa as a reflection of apartheid ideology and transformation.] In D. de La Fontaine, F. Müller, C. Hofmann, & B. Leubolt (Eds.), *Das politische System Südafrikas. [The political system of South Africa.]* (pp. 415–433). Springer VS.

Jürgens, U., Donaldson, R., Rule, S., & Bähr, J. (2013). Townships in South African cities. Literature review and research perspectives. *Habitat International*, *39*, 256–260.

Kauermann, G., & Carroll, R. J. (2001). A note on the efficancy of Sandwich Covariance Matrix Estimation. *Journal of the American Statistical Association, 96*(456), 1387–1396.

Kelle, U., & Kluge, S. (1999). *Vom Einzelfall zum Typus. Fallvergleich und Fallkontrastierung. [From individual case to type. Case comparison and case contrasting.]* Leske und Budrich.

Kelle, U., & Kluge, S. (2010). *Vom Einzelfall zum Typus. Fallvergleich und Fallkontrastierung in der qualitativen Sozialforschung. [From individual case to type. Case comparison and case contrasting in qualitative social research.]* Springer VS.

Kirchler, E., & Walenta, C. (2010). *Motivation. [Motivation.]* Facultas.

Kleinke, K., Schlüter, E., & Christ, O. (2017). *Strukturgleichungsmodelle mit Mplus. Eine praktische Einführung. [Structural equation models with Mplus. A practical introduction.]* De Gruyter.

Kline, R. B. (2012). Assumptions in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 111–125). Guilford Press.

Kline, R. B. (2016). *Principles and practice of structural equation modeling.* The Guilford Press.

Kloosterman, R., Notten, N., Tolsman, J., & Kraaykamp, G. (2011). The effect of parental reading socialization and early school involvement on children's academic performance. A panel study of primary school pupils in the Netherlands. *European Sociological Review, 27*(3), 291–306.

Kluge, S. (1999). *Empirisch begründete Typenbildung. Zur Konstruktion von Typen und Typologien in der qualitativen Sozialforschung. [Empirically based type construction. On the construction of types and typologies in qualitative social research.]* Leske und Budrich.

Kluge, S. (2000). *Empirically grounded construction of types and typologies in qualitative social research* (Forum: Qualitative Social Research Vol. 1, No. 1, Art. 14, Ed.). Retrieved August 30, 2022, from https://www.qualitative-research.net/index.php/fqs/article/view/1124/2500

Krog, A. (2000). *Country of my skull.* Three Rivers Press.

Kuhl, J. (1983). *Motivation, Konflikt und Handlungskontrolle. [Motivation, conflict, and action control.]* Springer.

Kühnel, S.-M. (1993). Lassen sich ordinale Daten mit linearen Strukturgleichungsmodellen analysieren? [Can ordinal data be analyzed with linear structural equation models?] *ZA-Information/Zentralarchiv für Empirische Sozialforschung*, *33*, 29–51.

La Guma, A. (Ed.). (1972). *Apartheid. A collection of writings on South African racism by South Africans.* Lawrence and Wishart.

Lam, D., Ardington, C., & Leibbrandt, M. (2011). Schooling as a lottery. Racial differences in school advancement in urban South Africa. *Journal of Development Economics*, *95*(2), 121–136.

Lamnek, S., & Krell, C. (2016). *Qualitative Sozialforschung. [Qualitative social research.]* Beltz.

Lareau, A. (1987). Social class differences in family-school relationships. The importance of cultural capital. *Sociology of Education*, *60*(2), 73–85.

Lareau, A. (2002). Invisible inequality. Social class and childrearing in Black families and White families. *American Sociological Review*, *67*(5), 747–776.

Larsen, R. (2011). Missing Data Imputation versus Full Information Maximum Likelihood with second-level dependencies. *Structural Equation Modeling: A Multidisciplinary Journal*, *18*(4), 649–662.

Lavenda, O. (2011). Parental involvement in school. A test of Hoover-Dempsey and Sandler's model among Jewish and Arab parents in Israel. *Children and Youth Services Review*, *33*(6), 927–935.

Lewin, K. (1926). *Vorsatz, Wille und Bedürfnis: Mit Vorbemerkungen über die psychischen Kräfte und Energien und die Struktur der Seele. [Intention, will and need: With preliminary remarks on the psychic powers and energies and the structure of the soul.]* Springer.

Li, C.-H. (2016). Confirmatory factor analysis with ordinal data. Comparing robust maximum likelihood and diagonally weighted least sqaures. *Behavior Research Methods*, *48*, 936–949.

Light, A., & Strayer, W. (2000). Determinants of Ccollege completion. School quality or student ability? *The Journal of Human Resources*, *35*(2), 299–332.

Little, R., & Rubin, D. B. (2020). *Statistical analysis with missing data*. Wiley.

Little, T. D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data. Practical and theoretical issues. *Multivariate Behavioral Research*, *32*(1), 53–76.

Locke, E. A. (1968). Toward a theory of task motivation and incentives. *Organizational Behavior and Human Performance*, *3*(2), 157–189.

Lüdtke, O., Robitzsch, A., Trautwein, U., & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. Probleme und Lösungen. [Dealing with missing values in psychological research. Problems and solutions.] *Psychologische Rundschau*, *58*(2), 103–117.

Maaz, K., Trautwein, U., Lüdtke, O., & Baumert, J. (2008). Educational transitions and differential learning enviornments. How explicitly between-school tracking contributes to social inequality in educational outcomes. *Child Development Persepectives*, *2*(2), 99–106.

Maccoby, E. E., & Martin, J. A. (1983). Socialization in the context of the family. Parent-child interaction. In E. M. Hetherington & P. H. Mussen (Eds.), *Handbook of child psychology: Vol. 4: Socialization, personality, and social development* (pp. 1–101). Wiley.

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules. Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal*, *11*(3), 320–341.

Martin, M. O., & Mullis, I. V. S. (2012a). *Methods and procedures in TIMSS and PIRLS 2011: prePIRLS 2011. Characteristics of national samples* (International Study

Center Lynch School of Education, Boston College, Ed.). IEA. Retrieved August 30, 2022, from https://timssandpirls.bc.edu/methods/pdf/preP11_National_Characteristics.pdf

Martin, M. O., & Mullis, I. V. S. (2012b). *Methods and procedures in TIMSS and PIRLS 2011: PIRLS 2011 Students Confident in Reading scale* (International Study Center Lynch School of Education, Boston College, Ed.). IEA. Retrieved August 30, 2022, from https://timssandpirls.bc.edu/methods/pdf/P11_R_Scales_SCR.pdf

Martin, M. O., & Mullis, I. V. S. (2012c). *Methods and procedures in TIMSS and PIRLS 2011: The PIRLS 2011 Students Like Reading scale* (International Study Center Lynch School of Education, Boston College, Ed.). IEA. Retrieved August 30, 2022, from https://timssandpirls.bc.edu/methods/pdf/P11_R_Scales_SLR.pdf

Martin, M. O., Mullis, I. V. S., Foy, P., & Arora, A. (2012). *Methods and procedures in TIMSS and PIRLS 2011: Creating and interpreting the TIMSS and PIRLS 2011 context questionnaire scales* (International Study Center Lynch School of Education, Boston College, Ed.). IEA. Retrieved August 30, 2022, from https://timssandpirls.bc.edu/methods/pdf/TP11_Context_Q_Scales.pdf

Marx, C. (2000). Folter und Rassismus. Südafrika während der Apartheid. [Torture and racism. South Africa during apartheid.] In P. Burschel, G. Distelrath, & S. Lembke (Eds.), *Das Quälen des Körpers. Eine historische Anthropologie der Folter. [Torturing the body. A historical anthropology of torture.]* (pp. 257–279). Böhlau Verlag.

Marx, C. (2017). Das Erbe der Apartheid. [The heritage of Apartheid.] In D. de La Fontaine, F. Müller, C. Hofmann, & B. Leubolt (Eds.), *Das politische System Südafrikas. [The political system of South Africa.]* (pp. 41–52). Springer VS.

Maslow, A. H. (1970). *Motivation and personality*. Harper & Row.

Mattingly, D. J., Prislin, R., McKenzie, T. L., Rodriguez, J. L., & Kayzar, B. (2002). Evaluating evaluations. The case of parent involvement programs. *Review of Educational Research*, *72*(4), 549–576.

MAXQDA. (2011). *Reference manual for the text analysis software MAX Qualitative-DataAnalysis 10* (MAXQDA, Ed.). Retrieved August 30, 2022, from https://www.maxqda.com/download/manuals/MAX10_manual_eng.pdf

Medjedovic, I. (2008). Sekundäranalyse qualitativer Interviewdaten. Problemkreise und offene Fragen einer neuen Forschungsstrategie. [Secondary analysis of qualitative interview data. Problem areas and open questions of a new research strategy.] *Historical Social Research*, *33*(3), 193–216.

Meredith, W. (1993). Measurment invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525–543.

Mestry, R. (2014). A critical analysis of the national norms and standards for school funding policy. Implications for soical justice and equity in South Africa. *Educational Management Administration and Leadership*, *42*(6), 851–867.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis. An expanded sourcebook*. Sage.

Miller, J. (n.d.). *South Africa*. Retrieved August 30, 2022, from https://www.behance.net/gallery/84738939/South-Africa

Millsap, R. E., & Olivera-Aguilar, M. (2012). Investigating measurement invariance using confirmatory factor analysis. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 380–392). Guilford Press.

Motala, S., & Sayeed, Y. (2009). *No fee schools in South Africa* (Consortium for Research on Educational Access, Transitions and Equity, Ed.) [Policy Brief]. Retrieved August 30, 2022, from https://assets.publishing.service.gov.uk/media/57a08b7ce5274a31e0000bae/Policy_Brief_7.pdf

Motshekga, A. (2012). *Address to the consultative conference of the National Association of School Governing Bodies (NASGBs)* (Republic of South Africa, Ed.). Retrieved August 30, 2022, from https://www.gov.za/address-consultative-conference-national-association-school-governing-bodies-nasgbs-mrs-angie

Mullis, I. V. S., Drucker, K. T., Preuschoff, C., Arora, A., & Stanco, G. M. (2012). *Methods and procedures in TIMSS and PIRLS 2011: Assessment framework and*

instrument development (International Study Center Lynch School of Education, Boston College, Ed.). IEA. Retrieved August 30, 2012, from https://timssandpirls.bc.edu/methods/pdf/tp_instrument_devel.pdfhttps://timssandpirls.bc.edu/methods/pdf/tp_instrument_devel.pdf

Mullis, I. V. S., Martin, M. O., Foy, P., & Drucker, K. T. (2012). *PIRLS 2011 International Results in Reading* (TIMSS and PIRLS International Study Center, Boston College, Ed.). IEA. Retrieved August 30, 2022, from https://timssandpirls.bc.edu/pirls2011/downloads/P11_IR_FullBook.pdf

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *PIRLS 2016 International Results in Reading* (TIMSS and PIRLS International Study Center, Boston College, Ed.). IEA. Retrieved August 30, 2022, from https://timssandpirls.bc.edu/pirls2016/international-results/

Mullis, I. V. S., Martin, M. O., Kennedy, A. M., Trong, K. L., & Sainsbury, M. (2009). *PIRLS 2011 Assessment Framework* (TIMSS and PIRLS International Study Center, Boston College, Ed.). IEA. Retrieved August 30, 2022, from https://timssandpirls.bc.edu/pirls2011/downloads/PIRLS2011_Framework.pdf

Mullis, I. V., Martin, M. O., Foy, P., & Hooper, M. (2015a). *TIMSS 2015 International Results in Mathematics* (TIMSS and PIRLS International Study Center, Boston College, Ed.). IEA. Retrieved August 30, 2022, from http://timssandpirls.bc.edu/timss2015/international-results/

Mullis, I. V., Martin, M. O., Foy, P., & Hooper, M. (2015b). *TIMSS 2015 International Results in Science* (TIMSS and PIRLS International Study Center, Boston College, Ed.). IEA. Retrieved August 30, 2022, from http://timssandpirls.bc.edu/timss2015/international-results/

Muthén, B. O. (2008). *Cronbach's alpha etc.* [Mplus Discussion/Confirmatory Factor Analysis]. Retrieved August 30, 2022, from https://www.statmodel.com/discussion/messages/9/213.html?1502318413

Muthén, B. O. (1998-2004). *Mplus Technical Appendices* (Muthén & Muthén, Ed.). Retrieved August 30, 2022, from https://www.statmodel.com/download/techappen.pdf

Muthén, B. O., Muthén, L. K., & Asparouhov, T. (2015). *Estimator choices with categorical outcomes*. Retrieved August 30, 2022, from www.statmodel.com/download/EstimatorChoices.pdf

Muthén, L. K. (2011). *Type = Complex* [Mplus Discussion/Multilevel Data/Complex Sample]. Retrieved August 30, 2022, from http://www.statmodel.com/discussion/messages/12/7975.html?1590863282

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus User's Guide* (Muthén & Muthén, Ed.) [Version 8]. Retrieved August 30, 2022, from https://www.statmodel.com/download/usersguide/MplusUserGuideVer_8.pdf

Nagengast, B., & Marsh, H. W. (2014). Motivation and engagement in science around the globe. Testing measurement invariance with multigroup structural equation models across 57 countries using PISA 2006. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of international large-scale assessment. Background, technical issues, and methods of data analysis* [Statistics in the Social and Behavioral Sciences Series] (pp. 317–344). Chapman & Hall/CRC.

Ndimande, B. (2016). School choice and inequalities in post-apartheid South Africa. *Global Education Review*, *3*(2), 33–49.

Nerdinger, F. W. (1995). *Motivation und Handeln in Organisationen. Eine Einführung. [Motivation and action in organizations. An introduction.]* Kohlhammer.

Okpala, C. O., Okpala, A. O., & Smith, F. E. (2001). Parental involvement, instructional expenditures, family socioeconomic attributes, and student achievement. *The Journal of Educational Research*, *95*(2), 110–115.

Patall, E. A., Cooper, H., & Robinson, J. C. (2008). Parent involvememt in homework. A research synthesis. *Review of Educational Research*, *78*(4), 1039–1101.

Patel, L., & Graham, L. (2012). How broad-based is broad-based black economic empowerment? *Development Southern Africa*, *29*(2), 193–207.

Peng, S. S., & Wright, D. (1994). Explanation of academic achievement of Asian American students. *The Journal of Educational Research*, *87*(6), 346–352.

Peraita, C., & Pastor, M. (2000). The primary school dropout in Spain. The influence of family background and labor market conditions. *Education Economics*, *8*(2), 157–168.

Platzky, L., & Walker, C. (1985). *The surplus people. Forced removals in South Africa.* Ravan Press.

Powell, J. A. (2012). *Racing to justice. Transforming our conceptions of self and other to build an inclusive society.* Indiana University Press.

Punter, R. A., Glas, C. A. W., & Meelissen, M. R. M. (2016). *Psychometric Framework for Modeling Parental Involvement and Reading Literacy.* Springer Open.

Raykov, T. (1997). Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research*, *32*(4), 329–353.

Raykov, T., & Marcoulides, G. A. (2019). Thanks coefficient alpha, we still need you! *Educational and Psychological Measurement*, *79*(1), 200–210.

Reddy, T. (2004). *Higher education and social transformation. South Africa Case Study.* Council on Higher Education.

Reiniger, T., & López, A. S. (2017). Parental involvement in municipal schools in Chile. Why do parents choose to get involved? *School Psychology International*, *38*(4), 363–379.

Republic of South Africa. (1996). Government Gazette No. 84 of 1996: South African Schools Act.

Republic of South Africa. (2004). Government Gazette No. 53 of 2003: Broad-Based Black Economic Empowerment Act.

Republic of South Africa. (2006). Government Gazette No. 869 of 2006: Amended National Norms and Standards for School Funding.

Republic of South Africa. (2012a). *Census 2011. Census in brief* (Republic of South Africa, Ed.). Statistics South Africa. Retrieved August 30, 2022, from https://

www.statssa.gov.za/census/census_2011/census_products/Census_2011_Census_in_brief.pdf

Republic of South Africa. (2012b). *Pocket Guide to South Africa 2011/2012: South Africa's people* (Republic of South Africa, Ed.). Government Communications and Information System. Retrieved August 30, 2022, from https://www.gcis.gov.za/sites/default/files/docs/resourcecentre/pocketguide/004_saspeople.pdf

Republic of South Africa. (2012c). *Pocket Guide to South Africa 2011/2012: History* (Republic of South Africa, Ed.). Government Communications and Information System. Retrieved August 30, 2022, from https://www.gcis.gov.za/sites/default/files/docs/resourcecentre/pocketguide/003_history.pdf

Republic of South Africa. (2012d). *Pocket Guide to South Africa 2011/2012: Health* (Republic of South Africa, Ed.). Government Communications and Information System. Retrieved August 30, 2022, from https://www.gcis.gov.za/sites/default/files/docs/resourcecentre/pocketguide/026_health.pdf

Republic of South Africa. (2012e). *Pocket Guide to South Africa 2011/2012: Education* (Republic of South Africa, Ed.). Government Communications and Information System. Retrieved August 30, 2022, from https://www.gcis.gov.za/sites/default/files/docs/resourcecentre/pocketguide/017_education.pdf

Republic of South Africa. (2012f). *The status of the language of learning and teaching (LOLT) in South African public schools. A quantitative overview* (Republic of South Africa, Ed.). Basic Education. Retrieved August 30, 2022, from https://www.education.gov.za/Portals/0/Documents/Reports/Status%20of%20LOLT.pdf?ver=2011-03-30-231358-000

Republic of South Africa. (2013). *Education Statistics in South Africa 2011* (Republic of South Africa, Ed.). Basic Education. Retrieved August 30, 2022, from https://www.education.gov.za/Portals/0/Documents/Reports/2011%20Education%20Statistics%20WEB.pdf?ver=2015-01-29-161525-400

Republic of South Africa. (2015). *Financial statistics of consolidated general government 2013/2014* [Statistical Release P9119.4]. Statistics South Africa. Retrieved August 30, 2022, from https://www.statssa.gov.za/publications/P91194/P911942014.pdf

Republic of South Africa. (2017). *Mortality and causes of death in South Africa 2017. Findings from death notification* (Republic of South Africa, Ed.) [Statistical Release P0309.3]. Statistics South Africa. Retrieved August 30, 2022, from https://www.statssa.gov.za/publications/P03093/P030932017.pdf

Republic of South Africa. (2018). *Quaterly Labour Force Survey. Quarter 1, 2018* (Republic of South Africa, Ed.) [Statistical Release P0211]. Statistics South Africa. Retrieved August 30, 2022, from https://www.statssa.gov.za/publications/P0211/P02111stQuarter2018.pdf

Republic of South Africa. (2021). *General Household Survey 2021* (Republic of South Africa, Ed.) [Statistical Release P0318]. Statistics South Africa. Retrieved August 30, 2022, from https://www.statssa.gov.za/publications/P0318/P03182021.pdf

Robbins, S. P., & Judge, T. A. (2017). *Organizational behavior*. Pearson.

Robitzsch, A. (2020). Why ordinal variables can (almost) always be treated as continuous variables. Clarifying assumptions of robust continuous and ordinal factor analysis estimation methods. *Frontiers in Education*, *5*(589965), 1–7.

Rogan, M. (2019). *Informal workers in urban South Africa. A statistical snapshot* (WIEGO, Ed.) [Statistical Brief No. 19]. Retrieved August 30, 2022, from https://www.wiego.org/sites/default/files/publications/files/Informal%20Workers%20in%20Urban%20South%20Africa%20SB%2019.pdf

Roithmayer, D. (2017). Access, adequacy and equality. The constitutionality of school fee financing in public education. *South African Journal on Human Rights*, *19*(3), 382–429.

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, *63*(3), 581–592.

Rutkowski, L., von Davier, M., & Rutkowski, D. (Eds.). (2014). *Handbook of international large-scale assessment. Background, technical issues, and methods of data analysis* [Statistics in the Social and Behavioral Sciences Series]. Chapman & Hall/CRC.

Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference Chi-Square test statistic. *Psychometrika*, *75*(2), 243–248.

Savalei, V. (2010). Expected versus observed information in SEM with incomplete normal and nonnormal data. *Psychological Methods*, *15*(4), 352–367.

Savalei, V., & Bentler, P. M. (2005). A statistically justified pairwise ML method for incomplete nonnormal data. A comparison with direct ML and pairwise ADF. *Structural Equation Modeling: A Multidisciplinary Journal*, *12*(2), 183–214.

Scherer, R. (2020). Analysing PIAAC data with structural equation modeling in Mplus. In D. Maehler & B. Rammstedt (Eds.), *Large-scale cognitive assessments. Analyzing PIAAC data* [Methodology of Educational Measurement and Assessment] (pp. 165–208). Springer.

Schreier, M., & Odağ, Ö. (2010). Mixed methods. In G. Mey & K. Mruck (Eds.), *Handbuch Qualitative Forschung in der Psychologie [Handbook of qualitative research in psychology]* (pp. 263–277). Springer VS.

Seekings, J., & Nattrass, N. (2005). *Class, race, and inequality in South Africa*. Yale University Press.

Sénéchal, M., & Young, L. (2008). The effect of family literacy interventions on children's acquisition of reading from kindergarten to grade 3. A meta-analytic review. *Review of Educational Research*, *78*(4), 880–907.

Seyfried, S. F., & Chung, I.-J. (2002). Parent involvement as parental monitoring of student motivation and parent expectations predicting later achievement among African American and European American middle school age students. *Journal of Ethnic and Cultural Diversity in Social Work*, *11*(1-2), 109–131.

Shai, L., Molefinyana, C., & Quinot, G. (2019). Public procurement in the context of Broad-Based Black Economic Empowerment (BBBEE) in South Africa. Lessons learned for sustainable public procurement. *Sustainability*, *11*(7164), 1–27.

Shisana, O., Labadarios, D., Rehle, T., Simbayi, L., Zuma, K., Dhansay, A., Reddy, P., Parker, W., Hoosain, E., Naidoo, P., Hongoro, C., Mchiza, Z., Steyn, N. P., Dwane, N., Makoae, M., Maluleke, T., Ramlagan, S., Zungu, N., Evans, M. G., ...

SANHANES-1 Team. (2013). *South African National Health and Nutrition Examination Survey (SANHANES-1)* (Republic of South Africa, Ed.). Health. Retrieved August 30, 2022, from https://repository.hsrc.ac.za/bitstream/handle/20.500.11910/2864/7844.pdf?sequence=1&isAllowed=y

Singh, K., Bickley, P. G., Keith, t. Z., Keith, P. B., Trivette, P., & Anderson, E. (1995). The efffects of four components of parental involvement on eigth-grade student achievement. Structural analysis of NELS-88 data. *School Psychology Review*, *24*(2), 299–317.

Singh, P., Mbokodi, S. M., & Msila, V. T. (2004). Black parental involvement in education. *South African Journal of Education*, *24*(4), 301–307.

Spaull, N. (2013). Poverty and privilege. Primary school inequality in South Africa. *International Journal of Educational Development*, *33*(5), 436–447.

Spaull, N. (2019). Equity. A price too high to pay? In N. Spaull & J. D. Jansen (Eds.), *South African schooling. The enigma of inequality. A study of the present situation and future possibilities* [Policy Implications of Research in Education] (pp. 1–24). Springer.

Stevenson, D. L., & Baker, D. P. (1987). The family-school relation and the child's school performance [Speical Issue: Schools and Development]. *Child Development*, *58*, 1348–1357.

Stocké, V. (2007a). Explaining educational decision and effects of families' social class position. An empirical test of the Breen–Goldthorpe model of educational attainment. *European Sociological Review*, *23*(4), 505–519.

Stocké, V. (2007b). *The motive for status maintenance and inequality in educational decisions. Which of the parents defines the reference point?* (Universitaet Mannheim, Ed.) [SonderForschungsBereich 504 (Rationalitätskonzepte, Entscheidungsverhalten und ökonomische Modellierung)]. Retrieved August 30, 2022, from https://madoc.bib.uni-mannheim.de/2547/1/dp07_20.pdf

Stocké, V. (2010). Der Beitrag der Theorie rationaler Entscheidung zur Erklärung von Bildungsungleichheit. [The contribution of rational choice theory to explaining educa-

tional inequality.] In G. Quenzel & K. Hurrelmann (Eds.), *Bildungsverlierer. Neue Ungleichheiten. [Educational Failures. New Inequalities.]* (pp. 73–94). Springer VS.

Sullivan, A. (2006). Students as rational decision-makers. The question of beliefs and attitudes. *London Review of Education*, *4*(3), 271–290.

Svetina, D., Rutkowski, L., & Rutkowski, D. (2020). Multiple-group invariance with categorical outcomes using updated guidelines. An illustration using Mplus and the lavaan/semTools packages. *Structural Equation Modeling: A Multidisciplinary Journal*, *27*(1), 111–130.

The Economist. (2017). South Africa has one of the world's worst education systems: Why it is bottom of the class. *The Economist*. Retrieved August 30, 2022, from https://www.economist.com/middle-east-and-africa/2017/01/07/south-africa-has-one-of-the-worlds-worst-education-systems

The World Bank. (2018). *Overcoming poverty and inequality in South Africa. An assessment of drivers, contraints and opportunities* (The World Bank, Ed.). Retrieved August 30, 2022, from https://documents1.worldbank.org/curated/en/530481521735906534/pdf/Overcoming-Poverty-and-Inequality-in-South-Africa-An-Assessment-of-Drivers-Constraints-and-Opportunities.pdf

The World Bank. (n.d.-a). *Data/Country Classification: World Bank country and lending groups*. Retrieved August 30, 2022, from https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

The World Bank. (n.d.-b). *Poverty and Inequality Platform: Poverty headcount ratio.* The World Bank. Retrieved August 30, 2022, from https://data.worldbank.org/indicator/SI.POV.DDAY

The World Bank. (n.d.-c). *Poverty and Inequality Platform: Country profile*. The World Bank. Retrieved August 30, 2022, from https://pip.worldbank.org/country-profiles/ZAF

Twele (formerly Radermacher), N. (2013). *Street Vendors in Cape Town. A case study on the educational needs of workers in Cape Town's informal sector* [Unpublished project report, DVV International Regional Office Southern Africa]. Cape Town.

Twele (formerly Radermacher), N. (2015). *Diskrepanzen zwischen Aufstiegsmotivationen- und Ressourcen. Eine Untersuchung des Investitionsverhaltens südafrikanischer Strassenhändler. [Discrepancies between upward motivations- and resources. An investigation of the investment behavior of South African street traders.]* [Unpublished master thesis, University of Bremen].

UNAIDS. (2021). *Country Factsheets: South Africa*. Retrieved August 30, 2022, from https://www.unaids.org/en/regionscountries/countries/southafrica

UNDP. (n.d.). *Human Development Reports: South Africa*. Retrieved August 30, 2022, from https://hdr.undp.org/data-center/specific-country-data#/countries/ZAF

UNDP. (2020). *Human Development Report 2020. The next frontier. Human development and the anthropocene* (UNDP, Ed.). Retrieved August 30, 2022, from https://hdr.undp.org/system/files/documents//hdr2020pdf.pdf

UNDP. (n.d.). *Human Development Reports: Human Development Index (HDI)*. Retrieved August 30, 2022, from https://hdr.undp.org/data-center/human-development-index#/indicies/HDI

UNESCO. (2012). *International Standard Classification of Eductaion: ISCED 2011* (UNESCO Institute for Statistics, Ed.). Retrieved August 30, 2022, from https://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf

UNESCO. (n.d.). *South Africa*. Retrieved August 30, 2022, from https://uis.unesco.org/en/country/za

United Nations. (n.d.). *Sustainable Development Goals*. Retrieved August 30, 2022, from https://sdgs.un.org/goals

Urban, D., & Mayerl, J. (2014). *Strukturgeleichungsmodellierung. Ein Ratgeber für die Praxis. [Structural equation modeling. A guide for practice.]* Springer.

van der Berg, S., Burger, C., Burger, R., de Vos, M., Du Rand, G., Gustafsson, M., Moses, E., Shepherd, D., Spaull, N., Taylor, S., van Broekhuizen, H., & von Fintel, D. (2011). *Low quality education as a poverty trap* (University of Stellenbosch, Department of Economics, Ed.) [Stellenbosch Economic Working Papers 25/11].

Retrieved August 30, 2022, from https://www.ekon.sun.ac.za/wpapers/2011/wp252011

van der Berg, S., Spaull, N., Wills, G., Gustafsson, M., & Kotzé, J. (2016). *Identifying binding constraints in education* (University of Stellenbosch, Department of Economics, Ed.) [Synthesis Report for the Programme to Support Pro-poor Policy Development (PSPPD)]. Retrieved August 30, 2022, from https://resep.sun.ac.za/wp-content/uploads/2017/10/PSPPD_BICiE-email-01062016.pdf

van Staden, S. (2010). *Reading between the lines. Contributing factors that affect grade 5 learner reading performance* [Doctoral dissertation, University of Pretoria]. Retrieved August 30, 2022, from https://repository.up.ac.za/handle/2263/24944

van Staden, S. (2020). Reading comprehension in early childhood. Putting new evidence of trend growth in South Africa into perspective. In K. S. Adeyemo (Ed.), *The education system of South Africa* [Global Education Systems] (pp. 1–18). Springer Cham.

van Staden, S., Bosker, R., & Bergbauer, A. (2016). Differences in achievement between home language and language of learning in South Africa. Evidence from prePIRLS 2011. *South African Journal of Childhood Education*, *6*(1), 1–10.

van Staden, S., & Howie, S. J. (2012). Reading between the lines. Contributing factors that affect Grade 5 student reading performance as measured across South Africa's 11 languages. *Educational Research and Evaluation*, *18*(1), 85–98.

van Staden, S., & Howie, S. J. (2014). Reflections on Creemers' Comprehensive Model of Educational Effectiveness for Reading Literacy. South African evidence from PIRLS 2006. *Perspectives in Education*, *32*(3), 172–192.

van Wyk, J. N. (1996). *Parental involvement in Black urban communities in Gauteng* [Doctoral dissertation, Univeristy of South Africa]. Retrieved August 30, 2022, from https://uir.unisa.ac.za/handle/10500/17781

Walker, J. M. T., Christa, L. I., Hoover–Dempsey, K. V., & Sandler, H. M. (2011). Latino parents' motivations for involvement in their children's schooling. An exploratory study. *The Elementary School Journal*, *111*(3), 409–429.

Walker, J. M. T., Wilkins, A. S., Dallaire, J. D., Sandlerand, H. M., & Hoover–Dempsey, K. V. (2005). Parental involvement. Model revision through scale development. *The Elementary School Journal*, *106*(2), 85–104.

Wang, X. (2013). Why students choose STEM majors. Motivation, high school learning, and postsecondary context of support. *American Educational Research Journal*, *50*(5), 1081–1121.

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables. Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling. Concepts, issues, and applications* (pp. 57–75). Sage.

Wigfield, A., & Eccles, J. S. (2000). Expectancy-Value Theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81.

Wigfield, A., Eccles, J. S., Simpkins, S., Roeser, R., & Schiefele, U. (2015). Development of achievement motivation and engagement. In R. M. Lerner, C. Garcia, & M. Lamb (Eds.), *Handbook of child psychology and developmental science: Vol. 3: Social and emotional development* (pp. 1–44). Wiley.

Wilder, S. (2014). Effects of parental involvement on academic achievement: a meta-synthesis. *Educational Review*, *66*(3), 377–397.

Wills, G. (2009). *South Africa's informal economy. A statistical profile* (WIEGO, Ed.) [WIEGO Working Paper (Urban Policies) No. 6]. Retrieved August 30, 2022, from https://www.wiego.org/sites/default/files/publications/files/Wills_WIEGO_WP6.pdf

Wirtz, M. A. (2004). Über das Problem fehlender Werte. Wie der Einfluss fehlender Analyseergebnisse entdeckt und reduziert werden kann. [About the problem of missing values. How to detect and reduce the impact of missing analysis results.] *Die Rehabilitation*, *43*(2), 109–115.

Witzel, A. (2000). *The problem-centered interview* (Forum: Qualitative Social Research Vol. 1, No. 1, Art. 22, Ed.). Retrieved August 30, 2022, from https://www.qualitative-research.net/index.php/fqs/article/view/1132/2522

Witzel, A., Medjedovic, I., & Kretzer, S. (2008). Sekundäranalyse qualitativer Daten. Zum Gegenwärtigen Stand neuer Forschungsstrategien. [Secondary analysis of qualitative data. On the Current Status of New Research Strategies.] *Historical Social Research*, *33*(3), 10–32.

Wolgast, A., Schwinger, M., Hahnel, C., & Stiensmeier-Pelster, J. (2017). Handling missing data in structural equation models in R. A replication study for applied researchers. *Electronic Journal of Research in Educational Psychology*, *15*(1), 5–47.

Xu, M., Kushner Benson, S. N., Mudrey-Camino, R., & Steiner, R. P. (2010). The relationship between parental involvement, self-regulated learning, and reading achievement of fifth graders. A path analysis using the ECLS-K database. *Social Psychology of Education*, *13*(2), 237–269.

Yang, G., Badri, M., Rashedi, A. A., & Almazroui, K. (2018). The role of reading motivation, self-efficacy, and home influence in students' literacy achievement. A preliminary examination of fourth graders in Abu Dhabi. *Large-Scale Assessments in Education*, *6*(10), 1–19.

Zimmerman, B. J. (2000). Self-efficacy. An essential motive to learn. *Contemporary Educational Psychology*, *25*(1), 82–91.

Zimmerman, L., Howie, S. J., & Smit, B. (2011). Time to go back to the drawing board. Organisation of primary school reading development in South Africa. *Educational Research and Evaluation*, *17*(4), 215–232.

A. Appendix: Qualitative Study

A.1. Interview Guideline

	Intention	Question
Introduction		Hello, welcome to today's interview! This interview is part of a case study by DVV International and The Big Issue (TBI) and deals with the question what kind of major educational needs and support workers of the informal sector wish for. Therefore we asked you to be a very important part of the intake. So I'd like to thank you for participating and sharing some details on the topic with me. Let me assure you, that all data will be handled carefully and anonymously. The interview will be split up in three parts, the past, the present and the future of your working and educational situation. Let's start in the past. You are a street vendor and selling The Big Issue in Cape Town ...
Former Situation	The intention of the category 'former situation' is to collect information related to the time before the interviewee became a street vendor and what motivated him to join the informal sector respectively become a street vendor. Especially for migrants, information on the reasons for moving to Cape Town is of interest.	What work did you do before you became a street vendor? Where did you live? Migrants: Why did you move to South Africa/Cape Town? What motivated you to become a street vendor? What advantages did working as a street vendor offer you?
Present Situation	The intention of the category 'present situation' is to collect information related to the current working situation. The interviewee is asked to reflect his or her situation by thinking about what and what not is satisfying about working as a street vendor. The second part asks the interviewee to outline his or her present qualifications for being a successful street vendor and to describe how he or she achieved those skills and knowledge. The third part is on a short evaluation about the current TBI social service network. The interviewees are asked to reflect the programme concerning the benefits it offers for his or her personal development and the chances to transfer to the formal job market.	What do you find or do not find satisfying about your current working situation? What kind of skills and knowledge do you need to be a successful TBI seller? How did you acquire those skills and knowledge? How does the social service network of TBI support you by being a successful street vendor? Which workshops and services do you find really helpful? What have you learned so far? How is the programme helpful or not helpful for your personal development? How is the programme helpful or not helpful in order to move over to the formal job market?
Future Situation	The intention of the category 'future situation' is to collect information concerning the future of the interviewee. He or she is asked to report on his or her wishes and goals concerning the future working situation. When the second part 'present situation' asked the interviewee to outline his or her current qualifications, the third part 'future situation' now asks the interviewee to reflect his or her lacks of skills and knowledge and outline what kind of further support and knowledge would be necessary to achieve personal goals.	What are your goals for the future concerning your working situation? If so, what would you like to do instead of being a TBI seller? Do you feel like your skills and education are enough to achieve these goals? If so what kind of support do you feel like you need in order to achieve these goals? What do you like to learn more about?
Conclusion		From your perspective, what advice would you give someone who is just starting to be a TBI seller after all? Is there any further information you'd like to share? Thank you for participating!

Table 11: Interview guideline for the Problem-Centered Interview. Own illustration.

A.2. Transcription Key

The interviews were recorded with the help of an audiotape recorder and transcribed using the transcription program F4. In order to avoid loss of information, all interviews were transcribed entirely. Transcripts were prepared by the interviewer herself. The following regularities were observed for the transcripts:

- Contributions of the interviewer during the interview are identified with 'I' (short for 'interviewer').
- Contributions of the interviewee are identified with 'IL' (short for 'interlocutor'). The abbreviation 'IL' is accompanied by the interviewee's spot on that day (i.e., _1 through _6, given that a total of 6 interviews was conducted on each of the two interview days) and the day that the interviewee's interview took place (i.e., day 1 or 2). For example, the first street vendor interviewed on day one is identified with the code 'IL01_1'.
- Transcripts document the interview word for word. Hence, informal language and colloquial language are used and false grammar is not corrected.
- In order to maintain anonymity of individuals, names mentioned during the interviews are made unrecognizable using [anonymized].
- Due to indistinct speech and accent, incomprehension of contributions occurred on the part of the interviewer as well as on the part of the interviewee. Such passages are identified with [-].
- In order to realistically capture the interview situation, completed statements are marked with '.' and statements that remained open are marked with '...' at the end of a sentence. Furthermore, commas have been discarded to avoid suggestions of indirect speech.
- Direct speech was marked with quotation marks (e.g., 'TEXT').
- Pauses and long pauses are marked with [pause] or [long pause] respectively.

- A laugh was marked with [laughing].
- TBI is an abbreviation for *The Big Issue*. The terms have been used frequently and interchangeably throughout the interviews.
- The South African currency is the South African rand, or simply rand (ZAR). In line with APA style, the official code ZAR is used to indicate the currency of a monetary value. In the case that the interviewee used the word 'rand', it was replaced with ZAR for coherency.

Due to the fact that the interviews took place in Cape Town/South Africa and therefore contain content that may be unfamiliar to the reader, footnotes with additional information were added at appropriate points.

A.3. Excerpts of Interview Statements

Types of Educational Investment Behavior	Selected Statements of Interviewees
Type I 'Trageted'	"Yes I tried to be participating in some workshops. I'm doing mosaic art" (IL01_1) "I'm learning the guitar at the moment" (IL01_1) "... especially the sales I mean the sales workshops they also add more ... [...] power to me you see" (IL05_1) "Yes so I'm gonna do grade 12 there at Philippi next year" (IL06_1) "I learned how to speak with my child" (IL06_1) "I went to Cape College you know I applied you know" (IL01_2) "Ehhm to put me futher up my study. I was in technical college of [-]. [...] I was being industrial electronics electric theory ehm math and all other subjects" (IL01_2
Type II 'Thwarted'	"But I believe we need [-] allowances we need allowances [-] so that you won't you won't be stuck. You won't starve by at least doing those learnerships. [...] Some people they don't have the time. They want to be working they want to be learning. [...] But now instead of working and learning. Which one would you choose? They obviously choose working where they'll be earning." (IL01_1) "We need some allowances attending those workshops. Because we don't earn that much" (IL01_1) "So that they can be learning while it's financially stable. [-] Because most of the people they don't learn while they won't be financially stable" (IL01_1) "Like a day allowance so that you can cope you be paid and [-] you'll be learning" (IL01_1) "[...] even if you give them a learnership they won't even find money that it takes to go those places" (IL01_1) "[...] the especially important one is the financial support. […] That is when I can concentrate on that course" (IL05_1) "But I haven't found the one that I will really like ehh feel like eh getting up and go for it" (IL01_2)
Type III ,Inconsistent'	"So I don't want to lie there have been so many offers but I haven't attended. There hasn't been the one that I will specifically like. [...] They do offer us like computer courses and all that. I don't want a computer course" (IL01_2) "But I haven't' found the one that I ehh feel like ehh getting up an go for it" (IL01_2)

Note:
If part of the statement has consciously been left out, [...] has been used to indicate that. Otherwise, the transcription key has been applied.

Table 12: Excerpts of interview statements collected from the complete interview transcripts. Own illustration, based on own analysis.

A.4. Techniques of Type Construction

1. **Development of relevant analysing dimensions**
   a. Research question, theoretical knowledge
   b. Sampling
   c. Guiding themes for the interviews
   d. Thematic coding of the interviews
   e. Dimensionalization, construction of variables ("quantification")
      1. (Thematic) case analysis
      2. (Thematic) case contrasts
2. **Grouping the cases and analysis of empirical regularities**
   a. Contrast of cases
   b. Computer-assisted grouping procedures (cluster analysis)
   c. Concept of attribute space
      1. Construction of the entire attribute space (cross tabulation)
      2. Assignment of the research elements
      3. Analysis of empirical regularities and – where appropriate – reduction of the attribute space
3. **Analysis of meaningful relationships and type construction**
   a. Consideration of further attributes
   b. Systematisation by computer-aided coding
   c. Search for "contradicting" and "deviating" cases
   d. Reduction of the attribute space to types
   e. Confrontation (of the individual cases with their ideal type)
4. **Characterization of the constructed types**
   a. Relevant analysing dimensions and further attributes
   b. Meaningful relationships
   c. Describing the typical through prototypes, ideal types, extreme types etc.

Figure 17: Techniques of type construction, applicable to the Model of Empirically Grounded Type Construction after (Kluge, 2000).

B. Appendix: Quantitative Results

B.1. Missings by Groups

Domain & Subdomain	Variable	Variable Description	% Missing, by Groups					
			Groups of HRL [Total % Missing *(% not administered)*]			Language Groups [Total % Missing *(% not administered)*]		
			LHRL	*HHRL*	*MHRL*	*Native*	*NonNative*	*Missing*
Parental Motivation:								
School Invitations for Involvement	ASBH10A (b, r)	Parents feel included in child's education	7.1 *(0)*	4.9 *(0)*	55.3 *(36.1)*	25.2 *(12.4)*	25.9 *(21.9)*	30.9 *(15.1)*
Educational Aspirations	ASBH18 (c)	Parental expectations of child's educational level	7.9 *(0)*	5.5 *(0)*	71.9 *(36.0)*	33.4 *(12.4)*	27.5 *(21.9)*	41.8 *(15.1)*
Student Outcomes in Learning:								
Academic Self-Efficacy	ASDGSCR (d)	Students feel confident in reading	15.7 *(0)*	8.5 *(0)*	17.0 *(0.7)*	15.6 *(0)*	4.6 *(0)*	24.8 *(1.8)*
Intrinsic Motivation	ASDGSLR (e)	Students like reading	6.0 *(0)*	3.1 *(0)*	8.3 *(0.7)*	6.0 *(0)*	1.4 *(0)*	14.1 *(1.8)*
Parental Involvement:								
Supervision	ASBH09A (a, r)	Parents discuss schoolwork	6.4 *(0)*	4.8 *(0)*	54.2 *(36.1)*	24.5 *(12.4)*	25.9 *(21.9)*	30.2 *(15.1)*
	ASBH09B (a, r)	Parents help with homework	8.2 *(0)*	5.3 *(0)*	55.2 *(36.1)*	26.0 *(12.4)*	25.8 *(21.9)*	31.5 *(15.1)*
	ASBH09C (a, r)	Parents ensure child sets aside time for homework	11.7 *(0)*	6.8 *(0)*	56.1 *(36.1)*	28.4 *(12.4)*	26.6 *(21.9)*	33.3 *(15.1)*
	ASBH09D (a, r)	Parents ask what child learned	10.1 *(0)*	6.3 *(0)*	56.0 *(36.1)*	27.6 *(12.4)*	26.3 *(21.9)*	32.1 *(15.1)*
	ASBH09E (a, r)	Parents check child's homework	10.8 *(0)*	6.5 *(0)*	56.1 *(36.1)*	28.0 *(12.4)*	26.6 *(21.9)*	32.5 *(15.1)*
Practice	ASBH09F (a, r)	Parents help practice reading	11.0 *(0)*	6.8 *(0)*	56.3 *(36.1)*	28.3 *(12.4)*	26.7 *(21.9)*	32.8 *(15.1)*
	ASBH09G (a, r)	Parents help practice math	11.1 *(0)*	6.3 *(0)*	56.2 *(36.1)*	28.1 *(12.4)*	26.6 *(21.9)*	33.0 *(15.1)*
	ASBH09H (a, r)	Parents talk with child about what child reads	10.6 *(0)*	6.5 *(0)*	55.8 *(36.1)*	27.7 *(12.4)*	26.5 *(21.9)*	32.7 *(15.1)*

Legend:
(r) Reverse coding has been applied to ensure concordant pairing of covariance matrices.
(a) Categories are: 1 'Never or almost never', 2 'Once or twice a month', 3 'Once or twice a week', 4 'Every day or almost every day'
(b) Categories are: 1 'Disagree a lot', 2 'Disagree a little', 3 'Agree a little', 4 'Agree a lot'
(c) Categories are: 1 'Grade 9/Standard 7', 2 'Grade 12/Standard 10', 3 'Post secondary training (Vocational training; e.g., College)', 4 'Technikon Diploma', 5 'First degree', 6 'Honours degree/Masters or PhD degree'
(d) Categories are: 1 'Confident', 2 'Somewhat confident', 3 'Not confident'
(e) Categories are: 1 'Like Reading', 2 'Somewhat like reading', 3 'Do not like reading'
(n.a.) Information is not meaningful.

Table 13: Rate of missingness, by groups. Unweighted results in percent. Own illustration, based on own analysis using the South African data set of prePIRLS 2011.

B.2. Factor Loadings (CFA)

STDYX Standardization

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
SUPVISE BY				
ASBH09AR	0.666	0.019	35.135	0.000
ASBH09BR	0.702	0.015	47.906	0.000
ASBH09CR	0.544	0.020	27.105	0.000
ASBH09DR	0.688	0.018	39.069	0.000
ASBH09ER	0.740	0.017	43.646	0.000
PRACTICE BY				
ASBH09FR	0.771	0.013	60.831	0.000
ASBH09GR	0.731	0.014	50.610	0.000
ASBH09HR	0.759	0.013	57.781	0.000

Table 14: Standardized factor loadings of the 2-factor solution obtained from CFA. Mplus output based on own analysis using the South African data set of prePIRLS 2011.

B.3. Chi² Contribution of Groups

Chi² Contribution by Groups						
	LHRL	*HHRL*	*LHRL*	*MHRL*	*Native*	*NonNative*
Baseline Model	**80.010**	**80.927**	**77.803**	**56.328**	**56.712**	**59.068**
Metric Invariance	88.303	96.555	87.451	69.946	63.050	69.835
Partial Metric Invariance	82.696	86.162	82.214	62.902	60.392	64.372
Scalar Invariance	111.439	120.087	97.799	82.189	122.030	136.482
Partial Scalar Invariance	91.772	101.798	91.223	73.844	68.610	79.087

Table 15: Chi² contribution of groups. Own illustration, based on own analysis using the South African data set of prePIRLS 2011.

C. Appendix: Summary of Results

C.1. Summary of Results (English)

Based on the psychological framework, that is the Rubicon Model of Action Phases after Heckhausen & Gollwitzer (1987), which allows for the investigation of separate phases of actional behavior, this study focuses educational investment behavior of socioeconomic low-performers in South Africa. On that note, this doctoral thesis addresses three key questions: (I) are socioeconomic low-performers in South Africa motivated to invest in education; (II) is the motivation transferred into (parental) educational investments; and (III) are parental educational investments beneficial to student outcomes in learning. Aimed at answering thereof, this study furthermore utilizes both, a qualitative and a quantitative perspective.

First, the qualitative study focuses general educational investment behavior among street vendors in Cape Town/South Africa. The data base for this study are qualitative interviews, conducted in an original study of 2013 on the educational needs of street vendors in Cape Town/South Africa. Aimed at analyzing thereof with regard to a secondary research question on the educational investment behavior of socioeconomic low-performers in South Africa, the Model of Empirically Grounded Type Construction after Kelle & Kluge (1999, 2010) was utilized. Overall, findings revealed three types of educational investment behavior among the focus-group. Types are type I *Targeted*, type II *Thwarted*, and type III *Inconsistent*. Whereas type II and III do not turn motivation into actions due to a lack of time and monetary resources, as well as the dilemma of *earning vs. learning*, type I acts on the motivation and takes educational investments. That is, for example, towards the improvement of skills and knowledge or even towards formal education. Thus, type I contradicts the assumptions drawn from Expectancy-Value Theory (EV-Theory) after Esser (1999). Used to substantiate the motivational phase indicated by the Rubicon Model, EV-Theory assumes that individuals of lower socioeconomic status are less likely to be motivated to take educational investments, and thus, refrain form taking them according to the outcomes of their cost-benefit calculations.

C APPENDIX: SUMMARY OF RESULTS

Motivated by the findings of the qualitative study, the quantitative study address educational investment behavior of the focus group from a different perspective. The study emphasizes parental involvement activities as a form of educational investment behavior, and explores its effects on student outcomes in learning. For that purpose, this study utilizes South African data of IEA's prePIRLS 2011, and applies structural equation models to mirror the phases of human action behavior as indicated by the Rubicon Model. A particularity of the models is the comparison of multiple groups of the South African population. For that purpose, this study utilizes comparison groups, counteracting the focus group. Groups are withdrawn from information on access to home resources for learning as well as language groups, given that those are characteristics considered to be decisive for the socioeconomic status of South African subpopulations. Results of quantitative analyses reveal that parental educational investment behavior among socioeconomic low-performers in South Africa are neither affected by a lack of motivation, nor the undertaking of actions with regard to supporting the child's education; but rather the quality thereof. Hence, results show significant effects between motivational factors and parental involvement activities. However, whereas parental involvement activities of the comparison group (i.e. socioeconomic high-performers) are significantly associated with student outcomes in learning, results did not show similar effects for the reference groups (i.e., socioeconomic low-performers).

To conclude, this study adds to the investigation of (parental) educational investment behavior, and more precisely, on the phases thereof, utilizing theories of human action behavior and rational decision making. On that note, results of this study reveal that the issue is not with the (parental) motivation for educational investments or to make an educational investment, but rather with the quality thereof. With regard to South Africa's political past and the long-term consequences of the Apartheid Regime, especially for the country's socioeconomic low-performers, the remaining question is whether the enhancement of the quality of parental educational investments holds the potential to make a difference for children's educational performance and careers, respectively. And thus, to help discontinuing the inheritance of social deprivation long-term. Results of this

study provide evidence for the assumption that this is true, and encourages policies to allow for further analyses on the subject.

No publications resulted from this doctoral thesis.

C.2. Summary of Results (German)

Auf der Grundlage eines psychologischen Rahmenmodells, nämlich des Rubikonmodells der Handlungsphasen nach Heckhausen & Gollwitzer (1987), welches die Untersuchung einzelner Phasen des menschlichen Handlungsverhaltens ermöglicht, konzentriert sich diese Studie auf das Bildungsinvestitionsverhalten von sozioökonomisch benachteiligten Personen in Südafrika. In diesem Sinne befasst sich diese Dissertation mit drei Schlüsselfragen: (I) Sind sozioökonomisch benachteiligte Personen in Südafrika motiviert, in Bildung zu investieren; (II) wird diese Motivation in Bildungsinvestitionen, auch im Sinne elterlicher Aktivitäten, umgesetzt; und (III) wirken sich elterliche Bildungsinvestitionen positiv auf den Bildungserfolg von Schüler:innen aus. Um diese Fragen zu beantworten, nimmt diese Arbeit sowohl eine qualitative als auch eine quantitative Perspektive ein.

Zunächst konzentriert sich die qualitative Studie auf das allgemeine Bildungsinvestitionsverhalten von Straßenhändler:innen in Kapstadt/Südafrika. Die Datengrundlage für diese Studie bilden qualitative Interviews, die im Rahmen der Originalstudie aus dem Jahr 2013 zum Bildungsbedarf von Straßenhändler:innen in Kapstadt/Südafrika durchgeführt wurden. Um diese im Hinblick auf eine sekundäre Forschungsfrage zum Bildungsinvestitionsverhalten sozioökonomisch benachteiligter Personen in Südafrika zu analysieren, wurde das Modell der empirisch begründeten Typenbildung nach Kelle & Kluge (1999, 2010) verwendet. Insgesamt ergeben die Ergebnisse drei Typen von Bildungsinvestitionsverhalten in der Fokusgruppe. Dabei handelt es sich um Typ I *Zielorientiert*, Typ II *Ausgebremst* und Typ III *Inkonsequent*. Während Typ II und III Bildungsmotivation nicht in Handlungen umsetzen, was zum Teil auf den Mangel an Zeit und finanziellen Ressourcen bzw. auf das Dilemma von *earning vs. learning* zurückzuführen ist, gelingt es Typ I Motivation in Bildungsinvestitionen zu übersetzen. Das heißt, zum Beispiel, mit Blick auf die Verbesserung von Fähigkeiten und Kenntnissen, oder sogar mit Blick auf die formale Bildung. Damit widerspricht Typ I den Annahmen der Werterwartungs-Theorie (EV-Theorie) nach Esser (1999). Die EV-Theorie, die zur Begründung der im Rubikon-Modell angedeuteten Motivationsphase herangezogen wird, geht davon aus, dass Individuen mit niedrigerem sozioökonomischen Status weniger motiviert sind, Bildungsinvestitionen zu tätigen. Sie

neigen eher dazu, diese entsprechend den Ergebnissen ihrer Kosten-Nutzen-Kalkulationen zu unterlassen.

Angeregt durch die Ergebnisse der qualitativen Studie, wird in der quantitativen Studie das Bildungsinvestitionsverhalten der Fokusgruppe aus einer anderen Perspektive betrachtet. Das heißt, die Studie hebt die elterlichen Beteiligungsaktivitäten als eine Form des Bildungsinvestitionsverhaltens hervor und untersucht deren Auswirkungen auf die Lernergebnisse von Schüler:innen. Zu diesem Zweck verwendet diese Studie südafrikanische Daten der IEA-Studie prePIRLS 2011 und wendet Strukturgleichungsmodelle an, um die Phasen des menschlichen Handlungsverhaltens gemäß dem Rubikon-Modell abzubilden. Eine Besonderheit dieser Studie ist der Vergleich mehrerer Gruppen der südafrikanischen Bevölkerung. Zu diesem Zweck werden in dieser Studie als Gegenstück zur Fokusgruppe Vergleichsgruppen eingesetzt. Die Gruppen wurden auf Basis von Informationen über den Zugang zu häuslichen Lernressourcen sowie Sprachkenntnissen gebildet, da dies Merkmale sind, die als entscheidend für den sozioökonomischen Status der südafrikanischen Bevölkerung gelten. Die Ergebnisse der quantitativen Analysen zeigen, dass das elterliche Bildungsinvestitionsverhalten von sozioökonomisch benachteiligten Eltern in Südafrika weder von mangelnder Motivation noch von der Durchführung von Maßnahmen zur Unterstützung der Bildung des Kindes beeinflusst wird, sondern vielmehr von deren Qualität. Die Ergebnisse zeigen signifikante Effekte zwischen Motivationsfaktoren und elterlichen Beteiligungsaktivitäten. Während das elterliche Engagement in der Vergleichsgruppe der sozioökonomisch besser gestellten Eltern signifikant mit den Lernergebnissen der Schüler zusammenhängt, zeigen die Ergebnisse für die Vergleichsgruppe der sozioökonomisch Benachteiligten keine vergleichbaren Effekte.

Zusammenfassend lässt sich sagen, dass diese Arbeit zur Untersuchung des (elterlichen) Bildungsinvestitionsverhaltens beiträgt. Genauer gesagt, zu dessen Phasen, indem sie Theorien des menschlichen Handlungsverhaltens und der rationalen Entscheidungsfindung nutzt. In diesem Sinne zeigen die Ergebnisse dieser Studie, dass es nicht um die (elterliche) Motivation für Bildungsinvestitionen geht, sondern um die Qualität dieser Investitionen. Im Hinblick auf die politische Vergangenheit Südafrikas und die langfristigen Folgen des

Apartheidregimes, insbesondere für die sozioökonomisch schwachen Bevölkerungsschichten des Landes, stellt sich die Frage, ob die Verbesserung der Qualität elterlicher Bildungsinvestitionen das Potenzial hat, sich positiv auf die Performanz bzw. die schulische Laufbahn der Kinder auszuwirken und somit langfristig dazu beizutragen, die Vererbung sozialer Benachteiligung einzudämmen. Die Ergebnisse dieser Studie liefern Belege für die Annahme, dass dies der Fall ist, und ermutigen somit die Politik, weitere Analysen zu diesem Thema zu ermöglichen.

Aus dieser Doktorarbeit gingen keine Publikationen hervor.

Affidavit

I certify that I have written this doctoral thesis on my own and that I have not used any auxiliary materials other than those indicated. All passages, which are taken from other texts in terms of wording or meaning, have been referenced with indication of the original sources (including the internet and other electronic text and data collections), and according to the standard rules of scientific citation. This also applies to illustrations, pictures, figures, tables and the like.

No commercial doctoral consulting service has been used. The work has not been accepted in another or in a previous doctoral program.

[The signature has been excluded for data protection reasons.]

Nadine Twele
Hamburg, September 2022

Effects of Educational Investments Within and Beyond a Framework of Home-Based Parental Involvement. Qualitative and Quantitative Analyses on South African Subpopulations.

Extended Appendix
of the
Doctoral Thesis

Nadine Twele

Hamburg, September 02, 2022

Contents

1 Short Questionnaire

Pre-Questionnaire
'Street Vendors in Cape Town. A case study on the educational needs of workers in Cape Town's informal sector'
December 2013
DVV International & The Big Issue Cape Town/South Africa

What is your age?

What is your gender?

- Male
- Female

What is the language that you primarily use at work?

What is the language that you primarily use at home?

What is the highest level of education you have completed?

- Primary education (6 years of school)
- Secondary education (12 years of school)
- University degree: _____
- No education completed

Others: _____

What is your current marital status?

- Single
- Married
- Divorced
- Widowed

Where were you born?

Do you currently have a permanent residence in Cape Town?

- Yes
- No

Where are you currently living in Cape Town?

How long have you been living there?

How many members are currently living in your household?

For how long have you been working for The Big Issue?

Do you have additional jobs to selling The Big Issue?

- Yes
- No

If yes, what kind of job(s)?

What is your current household income per week in South African Rand?

2 Interview Transcripts

[This part has been excluded for data protection reasons. For information on this part please contact the author on the following e-mail address: docthesis.nt.2023(at)gmail.com]

3 Interview Postscripts

[This part has been excluded for data protection reasons. For information on this part please contact the author on the following e-mail address: docthesis.nt.2023(at)gmail.com]

4 Overview of Statements

Placed Investments	
Positive	
Measures towards (further) education	My sister did try to put me in a college IL01_2 176 Ehhm to put me further up my study. I was in technical college of SA IL01_2 178 I was being industrial electronics electric theory ehm math and all other subjects IL01_2 180 Yeah you know that level. I went to Cape College you know I applied you know IL01_2 262 Yes I do register at Philippi IL06_1 166 Yes so I'm gonna do grade 12 there at Philippi next year IL06_1 168
Measures towards expansion of practical skills	Yes I tried to be participating in some workshops. I'm doing mosaic art IL01_1 60 I acquired in Zimbabwe but here I had to take another course also. That job preparedness IL01_1 54 I do attend workshops like where they train us about making sales and all that IL01_2 122 Yeah [-] the job job readiness ... to tell us how to do and how to sell ourselves and there is also this one for [pause] what do they call it like something like vocational training what they were offering just to go and do those industrial ... being trained in a few like ehh electrical carpentry IL02_1 112 Yeah in terms of workshops there is those workshops like parental guidance it did help me and art as well I'm improving in terms of my arts you see IL02_1 118 So I learn this things IL02_2 81 I attend all the workshops IL02_2 109

	Because each and every workshop I learn something. Each and every workshop I learn something IL02_2 117 For instance to safe the money. To be wise ehhh to be smart many things many many many things IL02_2 122-125 Yes. No matter in Wynberg in [-] I'm attending IL02_2 111 Yes I participate Well looking after the children how to treat the children Hmm and HIV abuse the other one IL03_1 89-94 I learn IL03_1 98 The workshop there is the workshop that I was making for the parent and child IL03_2 66 To care the child and the other ones the workshop for [pause] the other one is for for [-] on the street IL03_2 68 Yes I can learn from the ... How to care the child IL03_2 72-74 Even the other workshop for the people who's using the drugs I also used that one. I also learned trough that one IL03_2 82 So I also learned from this workshop IL03_2 86 Yes I'm also looking for the job of the security because I'm also trained the security course and I also have the cashiers course IL03_2 88 I got the certificate for that ones IL03_2 100 Ehm but that that is I went to a workshop once where they they were talking about skills. A lady came along and talked about how to market yourself and all this. I've been to so many of those in the past outside of the offices of The Big Issue you know in my early days ... IL04_2 328 ... In Johannesburg in Durban I've been to so many seminars and how to do this how to sell this

	IL04_2 330 Especially the sales I mean the sales workshops they also add more ... ... Power to me you see IL05_1 124-126 Yeah there were lots of business workshops also yeah I was there IL05_2 84 And then we we we learn a lot from that IL06_1 95 Because we didn't know other things. How to raise children IL06_1 97 Here's another workshop it was business IL06_1 99 Like the other workshop for it was for children IL06_1 91 I learned how to speak with my child IL06_1 105 Before I started I'm go to a workshop to The Big Issue I'm attend IL06_2 81 Yeah TB workshop I was attend yeah IL06_2 109 Yeah plenty workshop I attend at The Big Issue IL06_2 113 Like English hours I'm learn IL06_2 119 And I learning when to attract a customer you see. Yeah I'm learning like that IL06_2 124 Yeah I learn a lot of things IL06_2 114-115
Negative	
Dilemma Earning vs. Learning	But I believe we need [-] allowances we need allowances [-] so that you won't you won't be stuck you won't starve by at least doing those learnerships IL01_1 106 Like a day allowances so that you can cope you be paid and [-] you'll be learning IL01_1 110

	Some people they don't have the time. They want to be working they want to be learning IL01_1 114 But now instead of working and learning. Which one would you choose? They obviously choose working where they'll be earning IL01_1 116 So that they can be learning while it's financially stable. [-] Because most of the people they don't learn while they won't be financially stable IL01_1 192 Ahhahha even if you give them a learnership they won't even find money that it takes to go those places IL01_1 194 Yes they are helpful and ehm it's only that they need to be convenient. [-] We need some allowances attending those workshops. Because we don't earn that much you see IL01_1 82 Yes yes they will be after money for food also. [-] That's why I was in allowances IL01_1 196 Yeah you know that level. I went to Cape College you know I applied you know. But they wanted money upfront you see And I told as the person I am I can't pay those fees and all that you see If I can get that I will know me I can do my own thing you understand IL01_2 262-266 I like to be a nurse but it's difficult to put the money together and then to go to school IL02_2 43 Hmm the especially important one is the financial support. Yes financial support That is when I can concentrate on that course IL05_1 185-189 Yeah education I want to be back to learn Yeah I want but it's the money IL06_2 210-212
No measures towards educational investments	But I haven't found the one that I will really like ehh feel like eh getting up and go for it IL01_2 134 So I don't want to lie there have been so many offers but I haven't attended. There hasn't been the one that I will specifically like IL01_2 148

	They do offer us like computer courses and all that I don't want a computer course IL01_2 164 Well I don't use social services at all here IL04_2 120 I have not attend the workshops IL04_2 152
Planned Investments	
Attitudes towards investing in (further) education	
Positive	Learning is the most important thing IL01_1 114 Our learning doesn't end IL01_1 84 Yes learning is very important IL01_1 124 But eh for me wow all I need is education IL01_2 156 IL01/2: That's what I need education IL01_2 158 I want to learn more. It's not enough IL06_1 137 I do want to do the education because it's important to have the education IL06_1 162 Because there is nothing we can do without education IL06_1 164 Ready to learn yeah IL06_2 216
Attitudes towards formal (further) education	That is why I maybe I would like to like maybe get an M3 which is equivalent to matric IL01_2 258 Me I want to study for education IL01_2 166 You know maybe study agriculture course that will teach you about planting and all those things IL01_2 348 Or else study like electric thing you know IL01_2 352

	College go up to college and get level 3. Maybe at least if I can get level three I can further up myself IL01_2 354 Puh I need to learn to speak up in terms of hm to I mean to to upgrade my my my qualifications of education you see IL02_1 76 Ok at least I need a diploma IL02_1 162 To any institution [-] IL02_1 167-168 And some of us want to I mean to advance our education you see IL02_1 212 I need to go to school To learn yes my love IL02_2 176-179 So I was supposed to go back to grade 12. So I want to do my matric IL06_1 125 Then I want to to I want to take the nursing course IL06_1 129 I have to go to the college study IL06_1 143 But I want to be back to learn other thing IL06_2 214
Attitudes towards expansion of practical skills	I haven't got much skills. I need more skills IL01_1 106 I need to get more more skills So that I can move yes yes yes IL01_1 102-104 You can't just work without the knowledge you see. You have got to be improving your skills IL01_1 122 But I need I need some I need some like computers Yes yes hmhm IL01_1 134-136 Me I would like to be trained you know. Like as an electrician you see. Cause I know I've got a capability in that industry you understand IL01_2 174

	And living from organic food and all that. That I will love to do you know. If there was a course maybe in that besides doing electric course IL01_2 250 Exactly. As well to improve my art cause I also want to do some art lessons at The Big Issue IL02_1 72 Just like to have some courses like ehm any type of course [-] IL02_1 76 At least I need training skills. I mean you know yeah IL02_1 152 Ehh I need something like ehh [-] both theoretical and practical you see IL02_1 171-172 To open my company because The Big Issue is teaching me for a business you know to talk to the people to to understand the business IL02_2 75 Even my English I do not accept to be the English like this IL02_2 187 I like it yeah For cooking IL03_1 159-162 Yes I must training those course IL03_2 98 I I want to learn better English IL04_1 196 Yeah but eh I did keep on asking jobs especially the computer because my aim is on computer skills IL05_1 156 So the thing that I was expecting to get to did a computer course. From the base [-] especially the administration Just to work in an office. If I can get those certificates I can feel happy IL05_1 162-166 [-] Maybe like a pluming I must go to courses you see how much is it now for the courses? IL05_2 110 I want to learn about the how to speak with customers IL06_1 68

	Yeah before I want to learn anything kind of language IL06_2 168 And licenses. Yeah the driver license. You can't working with the tourists without license IL06_2 172-174
Negative	-

5 Data Preparation (SPSS Syntax)

***DATA PREPARATION.**

***GET DATA.**
*******Note: Data is the merged South African data of the international data base of prePIRLS 2011 (PIRLS and prePIRLS 2011 Assessment. Copyright © 2009 International Association for the Evaluation of Educational Achievement (IEA). Publisher: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College, available at https://timssandpirls.bc.edu/pirls2011/international-database.html).*

```
GET FILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_StudHome.sav".
```

*****RECODING.**

******DOMAIN: PARENTS' MOTIVATION FOR INVOLVEMENT.**
*******SUBDOMAIN: PERCEPTIONS OF GENERAL SCHOOL INVITATIONS.**

********ASBH10A: GEN\SCH\AGREE\INCLUDED IN CHILDS EDU.**

```
RECODE ASBH10A (1=4) (2=3) (3=2) (4=1) (MISSING = SYSMIS) (ELSE = COPY) INTO
ASBH10Ar.
VALUE LABELS ASBH10Ar 1 'DISAGREE A LOT' 2 'DISAGREE A LITTLE' 3 'AGREE A
LITTLE' 4 'AGREE A LOT'.
VARIABLE LABELS ASBH10Ar 'GEN\SCH\AGREE\INCLUDED IN CHILDS EDU'.
EXECUTE.

FREQUENCIES ASBH10A ASBH10Ar
/format=notable.
```

******DOMAIN: PARENTS' INVOLVEMENT FORMS.**
*******SUBDOMAIN: HOME-BASED BEHAVIORS.**

********ASBH09A:GEN\SCHWORK\OFTEN\DISCUSS SCHOOLWORK.**

```
RECODE ASBH09A (1=4) (2=3) (3=2) (4=1) (MISSING = SYSMIS) (ELSE = COPY) INTO
ASBH09Ar.
VALUE LABELS ASBH09Ar 1 'NEVER OR ALMOST NEVER' 2 'ONCE OR TWICE A
MONTH' 3 'ONCE OR TWICE A WEEK' 4 'EVERY DAY OR ALMOST EVERY DAY'.
VARIABLE LABELS ASBH09Ar 'HOW OFTEN DISCUSS SCHOOLWORK'.
EXECUTE.

FREQUENCIES ASBH09A ASBH09Ar
/format=notable.
```

********ASBH09B:GEN\SCHWORK\OFTEN\HELP WITH HOMEWORK.**

```
RECODE ASBH09B (1=4) (2=3) (3=2) (4=1) (MISSING = SYSMIS) (ELSE = COPY) INTO
ASBH09Br.
VALUE LABELS ASBH09Br 1 'NEVER OR ALMOST NEVER' 2 'ONCE OR TWICE A
MONTH' 3 'ONCE OR TWICE A WEEK' 4 'EVERY DAY OR ALMOST EVERY DAY'.
VARIABLE LABELS ASBH09Br 'HOW OFTEN HELP WITH HOMEWORK'.
EXECUTE.

FREQUENCIES ASBH09B ASBH09Br
```

/format=notable.

********ASBH09C:GEN\SCHWORK\OFTEN\TIME FOR HOMEWORK.**
RECODE ASBH09C (1=4) (2=3) (3=2) (4=1) (MISSING = SYSMIS) (ELSE = COPY) INTO ASBH09Cr.
VALUE LABELS ASBH09Cr 1 'NEVER OR ALMOST NEVER' 2 'ONCE OR TWICE A MONTH' 3 'ONCE OR TWICE A WEEK' 4 'EVERY DAY OR ALMOST EVERY DAY'.
VARIABLE LABELS ASBH09Cr 'HOW OFTEN TIME FOR HOMEWORK'.
EXECUTE.

FREQUENCIES ASBH09C ASBH09Cr
/format=notable.

********ASBH09D:GEN\SCHWORK\OFTEN\ASK WHAT LEARNED.**
RECODE ASBH09D (1=4) (2=3) (3=2) (4=1) (MISSING = SYSMIS) (ELSE = COPY) INTO ASBH09Dr.
VALUE LABELS ASBH09Dr 1 'NEVER OR ALMOST NEVER' 2 'ONCE OR TWICE A MONTH' 3 'ONCE OR TWICE A WEEK' 4 'EVERY DAY OR ALMOST EVERY DAY'.
VARIABLE LABELS ASBH09Dr 'HOW OFTEN ASK WHAT LEARNED'.
EXECUTE.

FREQUENCIES ASBH09D ASBH09Dr
/format=notable.

********ASBH09E:GEN\SCHWORK\OFTEN\CHECK CHILDS HOMEWORK.**
RECODE ASBH09E (1=4) (2=3) (3=2) (4=1) (MISSING = SYSMIS) (ELSE = COPY) INTO ASBH09Er.
VALUE LABELS ASBH09Er 1 'NEVER OR ALMOST NEVER' 2 'ONCE OR TWICE A MONTH' 3 'ONCE OR TWICE A WEEK' 4 'EVERY DAY OR ALMOST EVERY DAY'.
VARIABLE LABELS ASBH09Er 'HOW OFTEN CHECK CHILDS HOMEWORK'.
EXECUTE.

FREQUENCIES ASBH09E ASBH09Er
/format=notable.

********ASBH09F:GEN\SCHWORK\OFTEN\HELP PRACTICE READING.**
RECODE ASBH09F (1=4) (2=3) (3=2) (4=1) (MISSING = SYSMIS) (ELSE = COPY) INTO ASBH09Fr.
VALUE LABELS ASBH09Fr 1 'NEVER OR ALMOST NEVER' 2 'ONCE OR TWICE A MONTH' 3 'ONCE OR TWICE A WEEK' 4 'EVERY DAY OR ALMOST EVERY DAY'.
VARIABLE LABELS ASBH09Fr 'HOW OFTEN HELP PRACTICE READING'.
EXECUTE.

FREQUENCIES ASBH09F ASBH09Fr
/format=notable.

********ASBH09G:GEN\SCHWORK\OFTEN\HELP PRACTICE MATH.**
RECODE ASBH09G (1=4) (2=3) (3=2) (4=1) (MISSING = SYSMIS) (ELSE = COPY) INTO ASBH09Gr.
VALUE LABELS ASBH09Gr 1 'NEVER OR ALMOST NEVER' 2 'ONCE OR TWICE A MONTH' 3 'ONCE OR TWICE A WEEK' 4 'EVERY DAY OR ALMOST EVERY DAY'.
VARIABLE LABELS ASBH09Gr 'HOW OFTEN HELP PRACTICE MATH'.
EXECUTE.

```
FREQUENCIES ASBH09G ASBH09Gr
/format=notable.

******ASBH09H:GEN\SCHWORK\OFTEN\TALK ABOUT WHAT READS.
RECODE ASBH09H (1=4) (2=3) (3=2) (4=1) (MISSING = SYSMIS) (ELSE = COPY) INTO
ASBH09Hr.
VALUE LABELS ASBH09Hr 1 'NEVER OR ALMOST NEVER' 2 'ONCE OR TWICE A
MONTH' 3 'ONCE OR TWICE A WEEK' 4 'EVERY DAY OR ALMOST EVERY DAY'.
VARIABLE LABELS ASBH09Hr 'HOW OFTEN TALK ABOUT WHAT READS'.
EXECUTE.

FREQUENCIES ASBH09H ASBH09Hr
/format=notable.

****DOMAIN: STUDENT OUTCOMES IN LEARNING.
*****SUBDOMAIN: STUDENT INTRINSIC MOTIVATION TO READ.

******ASDGSLR: STUDENTS LIKE READING/IDX.
RECODE ASDGSLR (1=3) (2=2) (3=1) (MISSING = SYSMIS) (ELSE = COPY) INTO
ASDGSLRr.
VALUE LABELS ASDGSLRr 1 'DO NOT LIKE READING' 2 'SOMEWHAT LIKE READING' 3
'LIKE READING'.
VARIABLE LABELS ASDGSLRr 'STUDENTS LIKE READING/IDX'.
EXECUTE.

FREQUENCIES ASDGSLR ASDGSLRr
/format=notable.

******ASBGSLR: STUDENTS LIKE READING/SCL.
RECODE ASBGSLR (MISSING = SYSMIS) (ELSE = COPY) INTO ASBGSLRr.
VARIABLE LABELS ASBGSLRr 'STUDENTS LIKE READING/SCL'.
EXECUTE.

FREQUENCIES ASBGSLR ASBGSLRr
/format=notable.

*****SUBDOMAIN: STUDENT SELF-EFFICACY IN READING.

******ASDGSCR: STUDENTS CONFIDENT IN THEIR READING/IDX.
RECODE ASDGSCR (1=3) (2=2) (3=1) (MISSING = SYSMIS) (ELSE = COPY) INTO
ASDGSCRr.
VALUE LABELS ASDGSCRr 1 'NOT CONFIDENT' 2 'SOMEWHAT CONFIDENT' 3
'CONFIDENT'.
VARIABLE LABELS ASDGSCRr 'STUDENTS CONFIDENT IN THEIR READING/IDX'.
EXECUTE.

FREQUENCIES ASDGSCR ASDGSCRr
/format=notable.

******ASBGSCR: STUDENTS CONFIDENT IN THEIR READING/SCL.
RECODE ASBGSCR (MISSING = SYSMIS) (ELSE = COPY) INTO ASBGSCRr.
VARIABLE LABELS ASBGSCRr 'STUDENTS LIKE READING/SCL'.
EXECUTE.
```

```
FREQUENCIES ASBGSCR ASBGSCRr
/format=notable.
```

*****COMPUTING GROUPING VARIABLES.**

******LANGUAGE OF TESTING.**

*******Note: 1) Compute new variable ITLANGnew, combining language of testing (ITLANG, response categories are 11 official languages) with language of testing spoken at home (ASBG03; response categories are 1 'Always', 2 'Sometimes', 3 'Never'); 2) Recode ITLANGnew into grouping variable ITLANGnewr with three categroies.*

```
COMPUTE ITLANGnew = 99.

IF (ITLANG >=1 AND ASBG03 =3) ITLANGnew = 0.
IF (ITLANG = 1 AND ASBG03 <= 2) ITLANGnew = 1.
IF (ITLANG = 2 AND ASBG03 <= 2) ITLANGnew = 2.
IF (ITLANG = 3 AND ASBG03 <= 2) ITLANGnew = 3.
IF (ITLANG = 4 AND ASBG03 <= 2) ITLANGnew = 4.
IF (ITLANG = 5 AND ASBG03 <= 2) ITLANGnew = 5.
IF (ITLANG = 6 AND ASBG03 <= 2) ITLANGnew = 6.
IF (ITLANG = 7 AND ASBG03 <= 2) ITLANGnew = 7.
IF (ITLANG = 8 AND ASBG03 <= 2) ITLANGnew = 8.
IF (ITLANG = 9 AND ASBG03 <= 2) ITLANGnew = 9.
IF (ITLANG = 10 AND ASBG03 <= 2) ITLANGnew = 10.
IF (ITLANG = 11 AND ASBG03 <= 2) ITLANGnew = 11.

ADD VALUE LABELS ITLANGnew
0 'LoT not Language Spoken at Home'
1 'Afrikaans'
2 'English'
3 'isiNdebele'
4 'isiXhosa'
5 'isiZulu'
6 'Sepedi'
7 'Sesotho'
8 'Setswana'
9 'siSwati'
10 'Tshivenda'
11 'Xitsonga'
99 'Missing'.

MISSING VALUES ITLANGnew (99).
VARIABLE LABELS ITLANGnew 'Language of Test by Language Spoken at Home'.
EXECUTE.

FREQUENCIES ITLANG ASBG03 ITLANGnew.
EXECUTE.

RECODE ITLANGnew (0 = 0) (3 THRU 11=1) (1 THRU 2 =2) (MISSING = SYSMIS) INTO
ITLANGnewr.
VALUE LABELS ITLANGnewr 0 'LoT not LANGUGAE SPOKEN AT HOME' 1 'NATIVE
AFRICAN LANGUAGE' 2 'ENGLISH or AFRIKAANS'.
VARIABLE LABELS ITLANGnewr 'LANGUAGE OF TEST'.
```

```
EXECUTE.

FREQUENCIES ITLANGnew ITLANGnewr.
EXECUTE.
```

******HOME RESOURCES FOR LEARNING.**

*******Note: Recode scale on home resources for learning (ASBGHRL) into percentiles of 5/95, 10/90, 25/75 and 30/70 percent of cases of the distribution.*

```
******5/95 percentile.
RECODE ASBGHRL (LOWEST THRU 5.60408 = 1) (5.60409 THRU 10.66560 = 2)
(10.66561 thru 15.14335 = 3) (ELSE = 9) INTO HRL05.
VALUE LABELS HRL05 1 'Lowest 5%' 2 'Others' 3 'Highest 5%' 9 'Omitted or Invalid'.
VARIABLE LABELS HRL05 'HRL 5 Percentile'.
EXECUTE.

FREQUENCIES ASBGHRL
/format=notable.
FREQUENCIES HRL05.

******10/90 percentile.
RECODE ASBGHRL (LOWEST THRU 5.97964 = 1) (5.97965 THRU 10.1886 = 2) (10.1887
THRU 15.14335 = 3) (ELSE = 9) INTO HRL10.
VALUE LABELS HRL10 1 'Lowest 10%' 2 'Others' 3 'Highest 10%' 9 'Omitted or Invalid'.
VARIABLE LABELS HRL10 'HRL 10 Percentile'.
EXECUTE.

FREQUENCIES ASBGHRL
/format=notable.
FREQUENCIES HRL10.

******25/75 percentile.
RECODE ASBGHRL (LOWEST THRU 7.29889 = 1) (7.29890 THRU 9.22127 = 2) (9.22128
THRU 15.14335 = 3) (ELSE = 9) INTO HRL25.
VALUE LABELS HRL25 1 'Lowest 25%' 2 'Others' 3 'Highest 25%' 9 'Omitted or Invalid'.
VARIABLE LABELS HRL25 'HRL 25 Percentile'.
EXECUTE.

FREQUENCIES ASBGHRL
/format=notable.
FREQUENCIES HRL25.

******30/70 percentile.
RECODE ASBGHRL (LOWEST THRU 7.29889 = 1) (7.29890 THRU 8.8203 = 2) (8.8204
THRU 15.14335 = 3) (ELSE = 9) INTO HRL30.
VALUE LABELS HRL30 1 'Lowest 30%' 2 'Others' 3 'Highest 30%' 9 'Omitted or Invalid'.
VARIABLE LABELS HRL30 'HRL 30 Percentile'.
EXECUTE.

FREQUENCIES ASBGHRL
/format=notable.
FREQUENCIES HRL30.
```

*****CHECKING & DEFINING MISSING VALUES.**

```
FREQUENCIES
IDCNTRY IDSCHOOL IDCLASS IDSTUD
ASBH18
ASRREA01 ASRREA02 ASRREA03 ASRREA04 ASRREA05
TOTWGT JKZONE JKREP
/format=notable.

RECODE
ASBH18 (MISSING = SYSMIS) (ELSE = COPY) INTO ASBH18r.
EXECUTE.

FREQUENCIES ASBH18 ASBH18r.
```

***SAVING OUTFILE.**

```
OUTFILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_StudHome_Rec.sav".
```

*****SPLIT DATA BY PLAUSIBLE VALUES (PV).**

******PV1.**

```
GET FILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_StudHome_Rec.sav".
RENAME VARIABLES ASRREA01 = ASRREA.
EXECUTE.
DESCRIPTIVES ASRREA.

SAVE OUTFILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_PV1.sav"
/KEEP =
IDCNTRY IDSCHOOL IDCLASS IDSTUD ITLANGnewr
HRL05 HRL10 HRL25 HRL30
ASBH10Ar
ASBH18r
ASBH09Ar ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er ASBH09Fr ASBH09Gr ASBH09Hr
ASBGSLRr ASDGSLRr ASBGSCRr ASDGSCRr
ASRREA
TOTWGT JKZONE JKREP.
```

******PV2.**

```
GET FILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_StudHome_Rec.sav".
RENAME VARIABLES ASRREA02 = ASRREA.
EXECUTE.
DESCRIPTIVES ASRREA.

SAVE OUTFILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_PV2.sav"
/KEEP =
IDCNTRY IDSCHOOL IDCLASS IDSTUD ITLANGnewr
HRL05 HRL10 HRL25 HRL30
ASBH10Ar
ASBH18r
ASBH09Ar ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er ASBH09Fr ASBH09Gr ASBH09Hr
ASBGSLRr ASDGSLRr ASBGSCRr ASDGSCRr
```

```
ASRREA
TOTWGT JKZONE JKREP.

****PV3.
GET FILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_StudHome_Rec.sav".
RENAME VARIABLES ASRREA03 = ASRREA.
EXECUTE.
DESCRIPTIVES ASRREA.

SAVE OUTFILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_PV3.sav"
/KEEP =
IDCNTRY IDSCHOOL IDCLASS IDSTUD ITLANGnewr
HRL05 HRL10 HRL25 HRL30
ASBH10Ar
ASBH18r
ASBH09Ar ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er ASBH09Fr ASBH09Gr ASBH09Hr
ASBGSLRr ASDGSLRr ASBGSCRr ASDGSCRr
ASRREA
TOTWGT JKZONE JKREP.

****PV4.
GET FILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_StudHome_Rec.sav".
RENAME VARIABLES ASRREA04 = ASRREA.
EXECUTE.
DESCRIPTIVES ASRREA.

SAVE OUTFILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_PV4.sav"
/KEEP =
IDCNTRY IDSCHOOL IDCLASS IDSTUD ITLANGnewr
HRL05 HRL10 HRL25 HRL30
ASBH10Ar
ASBH18r
ASBH09Ar ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er ASBH09Fr ASBH09Gr ASBH09Hr
ASBGSLRr ASDGSLRr ASBGSCRr ASDGSCRr
ASRREA
TOTWGT JKZONE JKREP.

****PV5.
GET FILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_StudHome_Rec.sav".
RENAME VARIABLES ASRREA05 = ASRREA.
EXECUTE.
DESCRIPTIVES ASRREA.

SAVE OUTFILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_PV5.sav"
/KEEP =
IDCNTRY IDSCHOOL IDCLASS IDSTUD ITLANGnewr
HRL05 HRL10 HRL25 HRL30
ASBH10Ar
ASBH18r
ASBH09Ar ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er ASBH09Fr ASBH09Gr ASBH09Hr
ASBGSLRr ASDGSLRr ASBGSCRr ASDGSCRr
ASRREA
TOTWGT JKZONE JKREP.
```

*****SAVE DATA AS .dat.**

******PV1.**

```
GET FILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_PV1.sav".
RECODE All (SYSMIS= -9999).
EXECUTE.
FORMATS ALL (f15.2).
EXECUTE.
Set DECIMAL = DOT.
EXECUTE.
SORT CASES BY IDCNTRY IDSCHOOL IDCLASS IDSTUD.

WRITE OUTFILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_PV1.dat"
TABLE /all.
EXECUTE.
```

******PV2.**

```
GET FILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_PV2.sav".
RECODE All (SYSMIS= -9999).
EXECUTE.
FORMATS ALL (f15.2).
EXECUTE.
Set DECIMAL = DOT.
EXECUTE.
SORT CASES BY IDCNTRY IDSCHOOL IDCLASS IDSTUD.

WRITE OUTFILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_PV2.dat"
TABLE /all.
EXECUTE.
```

******PV3.**

```
GET FILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_PV3.sav".
RECODE All (SYSMIS= -9999).
EXECUTE.
FORMATS ALL (f15.2).
EXECUTE.
Set DECIMAL = DOT.
EXECUTE.
SORT CASES BY IDCNTRY IDSCHOOL IDCLASS IDSTUD.

WRITE OUTFILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_PV3.dat"
TABLE /all.
EXECUTE.
```

******PV4.**

```
GET FILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_PV4.sav".
RECODE All (SYSMIS= -9999).
EXECUTE.
FORMATS ALL (f15.2).
EXECUTE.
Set DECIMAL = DOT.
EXECUTE.
SORT CASES BY IDCNTRY IDSCHOOL IDCLASS IDSTUD.
```

```
WRITE OUTFILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_PV4.dat"
TABLE /all.
EXECUTE.
```

******PV5.**

```
GET FILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_PV5.sav".
RECODE All (SYSMIS= -9999).
EXECUTE.
FORMATS ALL (f15.2).
EXECUTE.
Set DECIMAL = DOT.
EXECUTE.
SORT CASES BY IDCNTRY IDSCHOOL IDCLASS IDSTUD.

WRITE OUTFILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_PV5.dat"
TABLE /all.
EXECUTE.
```

*****VARIABLES FOR IDENTIFICATION OF MISSINGNESS.**

******Note: Items with the extension 'om' distinguish between items being 'Omitted/Invalid' and items 'Not Administered' (i.e., 'System Missing').*

***GET DATA.**

```
GET FILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_StudHome_Rec.sav".
```

*****RECODING.**

******DOMAIN: PARENTS' MOTIVATION FOR INVOLVEMENT.**

```
RECODE ASBH10A (9 = 0) (1 THRU 4 = 2) (SYSMIS = 1) INTO IDX10Aom.
VALUE LABELS IDX10Aom 0 'OMITTED/INVALID' 1 'SYSTEM MISSING' 2 'VALID'.
VARIABLE LABELS IDX10Aom 'IDX IDENTIFICATION MISSING BY LEVELS'.
EXECUTE.

FREQUENCIES ASBH10A IDX10Aom.
EXECUTE.

RECODE ASBH18 (9 = 0) (1 THRU 5 = 2) (SYSMIS = 1) INTO IDX18om.
VALUE LABELS IDX18om 0 'OMITTED/INVALID' 1 'SYSTEM MISSING' 2 'VALID'.
VARIABLE LABELS IDX18om 'IDX IDENTIFICATION MISSING BY LEVELS'.
EXECUTE.

FREQUENCIES ASBH18 IDX18om.
EXECUTE.
```

******DOMAIN: PARENTS' INVOLVEMENT FORMS.**

```
RECODE ASBH09A (9 = 0) (1 THRU 4 = 2) (SYSMIS = 1) INTO IDX09Aom.
VALUE LABELS IDX09Aom 0 'OMITTED/INVALID' 1 'SYSTEM MISSING' 2 'VALID'.
VARIABLE LABELS IDX09Aom 'IDX IDENTIFICATION MISSING BY LEVELS'.
EXECUTE.
```

```
FREQUENCIES ASBH09A IDX09Aom.
EXECUTE.

RECODE ASBH09B (9 = 0) (1 THRU 4 = 2) (SYSMIS = 1) INTO IDX09Bom.
VALUE LABELS IDX09Bom 0 'OMITTED/INVALID' 1 'SYSTEM MISSING' 2 'VALID'.
VARIABLE LABELS IDX09Bom 'IDX IDENTIFICATION MISSING BY LEVELS'.
EXECUTE.

FREQUENCIES ASBH09B IDX09Bom.
EXECUTE.

RECODE ASBH09C (9 = 0) (1 THRU 4 = 2) (SYSMIS = 1) INTO IDX09Com.
VALUE LABELS IDX09Com 0 'OMITTED/INVALID' 1 'SYSTEM MISSING' 2 'VALID'.
VARIABLE LABELS IDX09Com 'IDX IDENTIFICATION MISSING BY LEVELS'.
EXECUTE.

FREQUENCIES ASBH09C IDX09Com.
EXECUTE.

RECODE ASBH09D (9 = 0) (1 THRU 4 = 2) (SYSMIS = 1) INTO IDX09Dom.
VALUE LABELS IDX09Dom 0 'OMITTED/INVALID' 1 'SYSTEM MISSING' 2 'VALID'.
VARIABLE LABELS IDX09Dom 'IDX IDENTIFICATION MISSING BY LEVELS'.
EXECUTE.

FREQUENCIES ASBH09D IDX09Dom.
EXECUTE.

RECODE ASBH09E (9 = 0) (1 THRU 4 = 2) (SYSMIS = 1) INTO IDX09Eom.
VALUE LABELS IDX09Eom 0 'OMITTED/INVALID' 1 'SYSTEM MISSING' 2 'VALID'.
VARIABLE LABELS IDX09Eom 'IDX IDENTIFICATION MISSING BY LEVELS'.
EXECUTE.

FREQUENCIES ASBH09E IDX09Eom.
EXECUTE.

RECODE ASBH09F (9 = 0) (1 THRU 4 = 2) (SYSMIS = 1) INTO IDX09Fom.
VALUE LABELS IDX09Fom 0 'OMITTED/INVALID' 1 'SYSTEM MISSING' 2 'VALID'.
VARIABLE LABELS IDX09Fom 'IDX IDENTIFICATION MISSING BY LEVELS'.
EXECUTE.

FREQUENCIES ASBH09F IDX09Fom.
EXECUTE.

RECODE ASBH09G (9 = 0) (1 THRU 4 = 2) (SYSMIS = 1) INTO IDX09Gom.
VALUE LABELS IDX09Gom 0 'OMITTED/INVALID' 1 'SYSTEM MISSING' 2 'VALID'.
VARIABLE LABELS IDX09Gom 'IDX IDENTIFICATION MISSING BY LEVELS'.
EXECUTE.

FREQUENCIES ASBH09G IDX09Gom.
EXECUTE.

RECODE ASBH09H (9 = 0) (1 THRU 4 = 2) (SYSMIS = 1) INTO IDX09Hom.
VALUE LABELS IDX09Hom 0 'OMITTED/INVALID' 1 'SYSTEM MISSING' 2 'VALID'.
```

```
VARIABLE LABELS IDX09Hom 'IDX IDENTIFICATION MISSING BY LEVELS'.
EXECUTE.

FREQUENCIES ASBH09H IDX09Hom.
EXECUTE.
```

******DOMAIN: STUDENT OUTCOMES IN LEARNING.**

```
RECODE ASDGSLR (9 = 0) (1 THRU 3 = 2) (SYSMIS = 1) INTO IDXSLRom.
VALUE LABELS IDXSLRom 0 'OMITTED/INVALID' 1 'SYSTEM MISSING' 2 'VALID'.
VARIABLE LABELS IDXSLRom 'IDX IDENTIFICATION MISSING BY LEVELS'.
EXECUTE.

FREQUENCIES ASDGSLR IDXSLRom.
EXECUTE.

RECODE ASDGSCR (9 = 0) (1 THRU 3 = 2) (SYSMIS = 1) INTO IDXSCRom.
VALUE LABELS IDXSCRom 0 'OMITTED/INVALID' 1 'SYSTEM MISSING' 2 'VALID'.
VARIABLE LABELS IDXSCRom 'IDX IDENTIFICATION MISSING BY LEVELS'.
EXECUTE.

FREQUENCIES ASDGSCR IDXSCRom.
EXECUTE.
```

***SAVING OUTFILE.**

```
OUTFILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_StudHome_Rec.sav".
```

6 Missing-Data-Diagnosis (SPSS Syntax)

```
* Encoding: UTF-8.
* Script created using the IEA IDB Analyzer (Version 4.0.39).

include file =
"C:\Users\UserName\AppData\Roaming\IEA\IDBAnalyzerV4\bin\Data\Templates\SPSS_Mac
ros\JB_PV.ieasps".

JB_PV infile="C:\Users\MergedData\PIRLSLit_2011_RSA_StudHome_Rec.sav"/
        cvar=IDCNTRY /
        almvars=IDX10AOM IDX18OM IDX09AOM IDX09BOM IDX09COM IDX09DOM
        IDX09EOM IDX09FOM IDX09GOM IDX09HOM IDXSLRom IDXSCRom /
        rootpv=ASRREA0 /
        tailpv=/
        npv=5/
        wgt=TOTWGT/
        nrwgt=150 /
        rwgt=/
        jkz=JKZONE/
        jkr=JKREP/
        jk2type=FULL/
        nomiss=Y/
        method=JRR/
        kfac=0/
        shrtcut=N/
        viewcod=N/
        ndec=2/
        clean = Y/
        strctry = N/
        intavg = Y/
        graphs=N/
        selcrit = /
        selvar = /
        outdir="C:\Users\MergedData\IDXom_ReaAchievement"/
        outfile="IDXOm_by_ReaAchievement".
```

7 Explanatory Factor Analysis (Mplus Syntax)

```
Mplus VERSION 8.1
MUTHEN & MUTHEN

INPUT INSTRUCTIONS

TITLE: EFA

DATA: FILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_PV1.dat";

VARIABLE: NAMES =
IDCNTRY IDSCHOOL IDCLASS IDSTUD ITLANGnewr
HRL05 HRL10 HRL25 HRL30
ASBH10Ar
ASBH18r
ASBH09Ar ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er
ASBH09Fr ASBH09Gr ASBH09Hr
ASBGSLRr ASDGSLRr ASBGSCRr ASDGSCRr
ASRREA
TOTWGT JKZONE JKREP;

USEVARIABLES =
ASBH09Ar
ASBH09Br
ASBH09Cr
ASBH09Dr
ASBH09Er
ASBH09Fr
ASBH09Gr
ASBH09Hr;

MISSING = ALL (-9999);
WEIGHT = TOTWGT;
STRATIFICATION = JKZONE;
CLUSTER = JKREP;

DEFINE: JKREP=JKREP+1000000*JKZONE;
!Creating unique cluster values for each stratum
!(http://www.statmodel.com/discussion/messages/12/361.html?1532563878)

ANALYSIS:
TYPE = COMPLEX EFA 1 3;
ESTIMATOR = MLR;

OUTPUT: TECH1 SAMPSTAT RESIDUAL;
```

8 Confirmatory Factor Analysis (Mplus Syntax)

```
Mplus VERSION 8.1
MUTHEN & MUTHEN

INPUT INSTRUCTIONS

TITLE: CFA_2-factor

DATA: FILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_PV1.dat";

VARIABLE: NAMES =
IDCNTRY IDSCHOOL IDCLASS IDSTUD ITLANGnewr
HRL05 HRL10 HRL25 HRL30
ASBH10Ar
ASBH18r
ASBH09Ar ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er
ASBH09Fr ASBH09Gr ASBH09Hr
ASBGSLRr ASDGSLRr ASBGSCRr ASDGSCRr
ASRREA
TOTWGT JKZONE JKREP;

USEVARIABLES =
ASBH09Ar
ASBH09Br
ASBH09Cr
ASBH09Dr
ASBH09Er
ASBH09Fr
ASBH09Gr
ASBH09Hr;

MISSING = ALL (-9999);
WEIGHT = TOTWGT;
STRATIFICATION = JKZONE;
CLUSTER = JKREP;

DEFINE: KREP=JKREP+1000000*JKZONE;
!Creating unique cluster values for each stratum
!(http://www.statmodel.com/discussion/messages/12/361.html?1532563878)

ANALYSIS:
TYPE = COMPLEX;
ESTIMATOR = MLR;

MODEL:
SupVise by ASBH09Ar ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er;
Practice by ASBH09Fr ASBH09Gr ASBH09Hr;

OUTPUT: TECH1 RESIDUAL STDYX;
```

9 Measurement Invariance Testing (Mplus Syntax)

9.1 Configural Invariance

*****Model on configural invariance for all pairs:**

```
Mplus VERSION 8.1
MUTHEN & MUTHEN

INPUT INSTRUCTIONS

TITLE: MI_Configural

DATA: FILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_PV1.dat";

VARIABLE: NAMES =
IDCNTRY IDSCHOOL IDCLASS IDSTUD ITLANGnewr
HRL05 HRL10 HRL25 HRL30
ASBH10Ar
ASBH18r
ASBH09Ar ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er
ASBH09Fr ASBH09Gr ASBH09Hr
ASBGSLRr ASDGSLRr ASBGSCRr ASDGSCRr
ASRREA
TOTWGT JKZONE JKREP;

USEVARIABLES =
ASBH09Ar
ASBH09Br
ASBH09Cr
ASBH09Dr
ASBH09Er
ASBH09Fr
ASBH09Gr
ASBH09Hr;

USEOBSERVATIONS = HRL25 EQ 1; !1 = LHRL
!USEOBSERVATIONS = HRL25 EQ 3; !3 = HHRL
!USEOBSERVATIONS = HRL25 EQ 9; !9 = MHRL
!USEOBSERVATIONS = ITLANGnewr EQ 1; !1 = Native
!USEOBSERVATIONS = ITLANGnewr EQ 2; !2 = NonNative

MISSING = ALL (-9999);
WEIGHT = TOTWGT;
STRATIFICATION = JKZONE;
CLUSTER = JKREP;

DEFINE: JKREP=JKREP+1000000*JKZONE;
!Creating unique cluster values for each stratum
!(http://www.statmodel.com/discussion/messages/12/361.html?1532563878)

ANALYSIS:
TYPE = COMPLEX;
```

```
ESTIMATOR = MLR;

MODEL:
SupVise by ASBH09Ar*
ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er;
SupVise@1;
Practice by ASBH09Fr* ASBH09Gr ASBH09Hr;
Practice@1;

OUTPUT: TECH1 SAMPSTAT RESIDUAL STDYX;
```

9.2 Baseline Model

*****Baseline model for all pairs:**

```
Mplus VERSION 8.1
MUTHEN & MUTHEN

INPUT INSTRUCTIONS

TITLE: MI_Baseline

DATA: FILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_PV1.dat";

VARIABLE: NAMES =
IDCNTRY IDSCHOOL IDCLASS IDSTUD ITLANGnewr
HRL05 HRL10 HRL25 HRL30
ASBH10Ar
ASBH18r
ASBH09Ar ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er
ASBH09Fr ASBH09Gr ASBH09Hr
ASBGSLRr ASDGSLRr ASBGSCRr ASDGSCRr
ASRREA
TOTWGT JKZONE JKREP;

USEVARIABLES =
ASBH09Ar
ASBH09Br
ASBH09Cr
ASBH09Dr
ASBH09Er
ASBH09Fr
ASBH09Gr
ASBH09Hr;

USEOBSERVATIONS = HRL25 EQ 1 OR HRL25 EQ 3;
GROUPING = HRL25 (1 = Low 3 = High);
!USEOBSERVATIONS = HRL25 EQ 1 OR HRL25 EQ 9;
!GROUPING = HRL25 (1 = Low 9 = Missed);
!USEOBSERVATIONS = ITLANGnewr EQ 1 OR ITLANGnewr EQ 2;
!GROUPING = ITLANGnewr (1 = Native 3 = NonNative);

MISSING = ALL (-9999);
```

```
WEIGHT = TOTWGT;
STRATIFICATION = JKZONE;
CLUSTER = JKREP;

DEFINE: JKREP=JKREP+1000000*JKZONE;
!Creating unique cluster values for each stratum
!(http://www.statmodel.com/discussion/messages/12/361.html?1532563878)

ANALYSIS: TYPE = COMPLEX;
ESTIMATOR = MLR;

MODEL:
SupVise by ASBH09Ar*
ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er;
SupVise@1;

Practice by ASBH09Fr*
ASBH09Gr ASBH09Hr;
Practice@1;

Model High:
!Model Missed:
!Model NonNative:

SupVise by ASBH09Ar
ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er;

Practice by ASBH09Fr ASBH09Gr ASBH09Hr;
[ASBH09Ar ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er];
[ASBH09Fr ASBH09Gr ASBH09Hr];

[SupVise-Practice@0];

OUTPUT: TECH1 SAMPSTAT RESIDUAL STDYX;
```

9.3 (Partial) Metric Invariance

*****Model on (partial) metric invariance for all pairs:**

```
Mplus VERSION 8.1
MUTHEN & MUTHEN

INPUT INSTRUCTIONS

TITLE: MI_Metric

DATA: FILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_PV1.dat";

VARIABLE: NAMES =
IDCNTRY IDSCHOOL IDCLASS IDSTUD ITLANGnewr
HRL05 HRL10 HRL25 HRL30
ASBH10Ar
ASBH18r
```

```
ASBH09Ar ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er
ASBH09Fr ASBH09Gr ASBH09Hr
ASBGSLRr ASDGSLRr ASBGSCRr ASDGSCRr
ASRREA
TOTWGT JKZONE JKREP;

USEVARIABLES =
ASBH09Ar
ASBH09Br
ASBH09Cr
ASBH09Dr
ASBH09Er
ASBH09Fr
ASBH09Gr
ASBH09Hr;

USEOBSERVATIONS = HRL25 EQ 1 OR HRL25 EQ 3;
GROUPING = HRL25 (1 = Low 3 = High);
!USEOBSERVATIONS = HRL25 EQ 1 OR HRL25 EQ 9;
!GROUPING = HRL25 (1 = Low 9 = Missed);
!USEOBSERVATIONS = ITLANGnewr EQ 1 OR ITLANGnewr EQ 2;
!GROUPING = ITLANGnewr (1 = Native 3 = NonNative);

MISSING = ALL (-9999);
WEIGHT = TOTWGT;
STRATIFICATION = JKZONE;
CLUSTER = JKREP;

DEFINE: JKREP=JKREP+1000000*JKZONE;
!Creating unique cluster values for each stratum
!(http://www.statmodel.com/discussion/messages/12/361.html?1532563878)

ANALYSIS:
TYPE = COMPLEX;
ESTIMATOR = MLR;

MODEL:
SupVise by ASBH09Ar*
ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er;
SupVise@1;

Practice by ASBH09Fr* ASBH09Gr ASBH09Hr;
Practice@1;

Model High:

!Partial Metric Invariance:
!SupVise by ASBH09Br;
!Practice by ASBH09Hr;

[ASBH09Ar ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er];
[ASBH09Fr ASBH09Gr ASBH09Hr];

[SupVise-Practice@0];
SupVise-Practice*;
```

```
!Model Missed:

!Partial Metric Invariance:
!SupVise by ASBH09Dr;

![ASBH09Ar ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er];
![ASBH09Fr ASBH09Gr ASBH09Hr];

![SupVise-Practice@0];
!SupVise-Practice*;

!Model NonNative:

!Partial Metric Invariance:
!SupVise by ASBH09Dr;
![ASBH09Ar ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er];
![ASBH09Fr ASBH09Gr ASBH09Hr];
![SupVise-Practice@0];
!SupVise-Practice*;

OUTPUT: TECH1 SAMPSTAT RESIDUAL STDYX MODINDICES (4);
```

9.4 (Partial) Scalar Invariance

*****Model on (partial) scalar invariance for all pairs:**

```
Mplus VERSION 8.1
MUTHEN & MUTHEN

INPUT INSTRUCTIONS

TITLE: MI_Scalar

DATA: FILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_PV1.dat";

VARIABLE: NAMES =
IDCNTRY IDSCHOOL IDCLASS IDSTUD ITLANGnewr
HRL05 HRL10 HRL25 HRL30
ASBH10Ar
ASBH18r
ASBH09Ar ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er
ASBH09Fr ASBH09Gr ASBH09Hr
ASBGSLRr ASDGSLRr ASBGSCRr ASDGSCRr
ASRREA
TOTWGT JKZONE JKREP;

USEVARIABLES =
ASBH09Ar
ASBH09Br
ASBH09Cr
ASBH09Dr
ASBH09Er
ASBH09Fr
ASBH09Gr
ASBH09Hr;
```

```
USEOBSERVATIONS = HRL25 EQ 1 OR HRL25 EQ 3;
GROUPING = HRL25 (1 = Low 3 = High);
!USEOBSERVATIONS = HRL25 EQ 1 OR HRL25 EQ 9;
!GROUPING = HRL25 (1 = Low 9 = Missed);
!USEOBSERVATIONS = ITLANGnewr EQ 1 OR ITLANGnewr EQ 2;
!GROUPING = ITLANGnewr (1 = Native 3 = NonNative);

MISSING = ALL (-9999);
WEIGHT = TOTWGT;
STRATIFICATION = JKZONE;
CLUSTER = JKREP;

DEFINE: JKREP=JKREP+1000000*JKZONE;
!Creating unique cluster values for each stratum
!(http://www.statmodel.com/discussion/messages/12/361.html?1532563878)

ANALYSIS:
TYPE = COMPLEX;
ESTIMATOR = MLR;

MODEL:
SupVise by ASBH09Ar*
ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er;
SupVise@1;

Practice by ASBH09Fr* ASBH09Gr ASBH09Hr;
Practice@1;

Model High:

!Partial Scalar Invariance:
![ASBH09Cr ASBH09Dr];

SupVise-Practice*;

!Model Missed:

!Partial Scalar Invariance:
![ASBH09Ar ASBH09Dr];

!SupVise-Practice*;

!Model NonNative:

!Partial Scalar Invariance:
![ASBH09Br ASBH09Cr];

!SupVise-Practice*;
```

10 Structural Equation Model (Mplus Syntax)

*****Unrestricted (U) and Restricted (R) model for all pairs:**
******Note: Data for this model is .txt, including paths to five data sets, containing one PV each*

```
Mplus VERSION 8.1
MUTHEN & MUTHEN

INPUT INSTRUCTIONS

TITLE: SEM

DATA: FILE = "C:\Users\MergedData\PIRLSLit_2011_RSA_AllPVs.txt";
TYPE = IMPUTATION;

VARIABLE: NAMES =
IDCNTRY IDSCHOOL IDCLASS IDSTUD ITLANGnewr
HRL05 HRL10 HRL25 HRL30
ASBH10Ar
ASBH18r
ASBH09Ar ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er
ASBH09Fr ASBH09Gr ASBH09Hr
ASBGSLRr ASDGSLRr ASBGSCRr ASDGSCRr
ASRREA
TOTWGT JKZONE JKREP;

USEVARIABLES =
ASBH10Ar
ASBH18r
ASBH09Ar
ASBH09Br
ASBH09Cr
ASBH09Dr
ASBH09Er
ASBH09Fr
ASBH09Gr
ASBH09Hr
ASBGSLRr
ASBGSCRr
ASRREA;

USEOBSERVATIONS = HRL25 EQ 1 OR HRL25 EQ 3;
GROUPING = HRL25 (1 = Low 3 = High);
!USEOBSERVATIONS = HRL25 EQ 1 OR HRL25 EQ 9;
!GROUPING = HRL25 (1 = Low 9 = Missed);
!USEOBSERVATIONS = ITLANGnewr EQ 1 OR ITLANGnewr EQ 2;
!GROUPING = ITLANGnewr (1 = Native 3 = NonNative);

MISSING = ALL (-9999);
WEIGHT = TOTWGT;
STRATIFICATION = JKZONE;
CLUSTER = JKREP;

DEFINE: JKREP=JKREP+1000000*JKZONE;
!Creating unique cluster values for each stratum
```

!(http://www.statmodel.com/discussion/messages/12/361.html?1532563878)

```
ANALYSIS:
TYPE = COMPLEX;
ESTIMATOR = MLR;

MODEL:
SupVise by ASBH09Ar*
ASBH09Br ASBH09Cr ASBH09Dr ASBH09Er;
SupVise@1;

Practice by ASBH09Fr* ASBH09Gr ASBH09Hr;
Practice@1;

SupVise on ASBH10Ar (U01);
SupVise on ASBH18r (U02);
Practice on ASBH10Ar (U03);
Practice on ASBH18r (U04);
ASRREA on SupVise (U05);
ASRREA on Practice (U06);
ASBGSCRr on SupVise (U07);
ASBGSCRr on Practice (U08);
ASBGSLRr on SupVise (U09);
ASBGSLRr on Practice (U10);
SupVise with Practice (U11);
ASRREA with ASBGSCRr (U12);
ASRREA with ASBGSLRr (U13);
ASBGSCRr with ASBGSLRr (U14);

Model High:

SupVise by ASBH09Br;
Practice by ASBH09Hr;

!Model Missed:

!SupVise by ASBH09Dr;

!Model NonNative:

!SupVise by ASBH09Dr;

SupVise-Practice*;

SupVise on ASBH10Ar (R01);
SupVise on ASBH18r (R02);
Practice on ASBH10Ar (R03);
Practice on ASBH18r (R04);
ASRREA on SupVise (R05);
ASRREA on Practice (R06);
ASBGSCRr on SupVise (R07);
ASBGSCRr on Practice (R08);
ASBGSLRr on SupVise (R09);
ASBGSLRr on Practice (R10);
SupVise with Practice (R11);
ASRREA with ASBGSCRr (R12);
ASRREA with ASBGSLRr (R13);
```

```
ASBGSCRr with ASBGSLRr (R14);

!Implementing Equality constraints for restricted SEM:
MODEL CONSTRAINT:
!U01 = R01;
!U02 = R02;
!U03 = R03;
!U04 = R04;
!U05 = R05;
!U06 = R06;
!U07 = R07;
!U08 = R08;
!U09 = R09;
!U10 = R10;
!U11 = R11;
!U12 = R12;
!U13 = R13;
!U14 = R14;

OUTPUT: STDYX RESIDUAL TECH1;
```