The Econometrics of Hope

by

Nargiza Chorieva

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Agricultural and Resource Economics

Department of Resource Economics and Environmental Sociology
University of Alberta

Abstract

We rarely think of hope as a determinant of economic outcomes. Yet it has been shown to be a motivator of human behavior in the fields of psychology and anthropology. Recent studies have incorporated the longstanding psychological concept of hope into economics, but have failed to pay sufficient attention to measurement and statistical considerations that are required to operationalize hope within empirical econometric frameworks. This study applies advanced measurement techniques on hope survey data to derive hope measures that are meaningful for economic analysis. Our first objective is to ascertain the efficacy of standard survey measures of hope used commonly in the psychology literature. Our second objective is to assess possible sources of heterogeneity between individuals with different demographic characteristics and access to public services. To meet our objectives, we rely on a set of primary data collected in two rural regions of Tanzania with a wealth of mindset questions (e.g., hope, aspirations, self-efficacy, social cohesion). We address our first objective using Item Response Theory (IRT) methods. We find clear evidence that respondents perceive positive and negative questions differently. The former performs well at discriminating people with low levels of hope, while the latter are useful for identifying individuals with high levels of hope. With regards to our second objective, which we address using Differential Item Functioning (DIF) methods, we find significant evidence of heterogeneity. For example, we found high DIF in the responses of adults and adolescents. For instance, if DIF is ignored, adolescents will be erroneously scored as less hopeful than adults. These findings help us understand psychological differences that underpin subgroup responses and allow us to consider

them while making causal analyses.

To my Mom and Dad

Acknowledgements

I would like to thank my supervisors, Dr. Brent Swallow and Dr. Sandeep Mohapatra, for your patience, support, and belief in my ability to get it done. Thank you for the opportunity to dive into the field of development economics with a flavor of psychology. I learned from you to care for others, always be intellectually curious, think bigger and how to love your work.

I would also like to acknowledge the generous funding from World Vision who made this research possible. My appreciation goes to all the members of World Vision International and Tanzania teams for their contribution in making people's lives better. Thanks to Ifakara Health Institute for all that they contributed to the study design, data collection and management, ethics approvals, and feedback.

I also want to thank my friends in Edmonton, Astana, and Tashkent for your support during my hard times. Finally, let me thank Adilbek for always being by my side. I can't thank you enough.

Big thanks to my Mom and Dad. It wouldn't have been possible without you. Mom, you are always in my heart.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

In recent years, the psychological construct of hope has received increasing attention from economists, leading to the emergence of a new field that blends psychology, economics, and anthropology. Empirical evidence shows that hope and aspirations affect economic behavior such as individuals' motivations for savings and investments in their futures through entrepreneurship, education, health, or working with others to solve social issues (Banerjee et al. 2015; Duflo 2012; Lybbert and Wydick 2018; Lybbert et al. 2016). The economic approach to the theory of hope considers how our attitudes toward the future affect our current behavior. According to the emerging literature on the economics of hope (Lybbert and Wydick 2018), an individual's psychological attributes, including hope and aspiration levels create "internal constraints" that play a significant role in welfare and poverty outcomes. This contrasts with standard approaches in economics which are typically more focused on how to relieve individuals' "external constraints" such as a lack of access to schools, hospitals, roads, and markets (Besley 2017; Glewwe, Ross, and Wydick 2014).

A number of studies have demonstrated that additional interventions addressing such internal constraints significantly complement the results obtained from interventions that exclusively target external constraints (Banerjee et al. 2015; Bloem and Wydick 2021). One example is a study conducted in the Philippines where children who attended a faith-based kindergarten program had a higher level of academic

performance compared to children in a government-sponsored kindergarten program (Bloem and Wydick 2021). According to the authors, the difference may be explained by the mediating effects of socio-emotional skills such as self-control, openness, and conscientiousness among the faith-based kindergarten group. Thus, in this example an intervention targeting internal constraints, i.e., the faith-based program, significantly complemented the intervention addressing the external constraints, i.e., the standard education program. Although the study participants were not randomly assigned, Bloem and Wydick complemented the results of a randomized control study by Bryan, Choi, and Karlan (2020) who found that religious education increased income in ultra-poor Filipino households, possibly by increasing grit.

At the same time, recent field experiments show that attitudes towards hopelessness and risk aversion pose obstacles to growth and human development in many developing countries. Duflo (2012) argues that providing grounds for hope helps break the "poverty trap" - a set of self-reinforcing mechanisms where current poverty is a direct cause of future poverty. Duflo (2012) argues that many of the very poor may avoid investing their limited resources in their futures because they believe such efforts are unlikely to yield significant benefits. This belief can make it rational for those living in extreme poverty to constantly prioritize immediate consumption above future-oriented behaviors like education, which could help them get ahead and out of poverty. A lack of hope for a better life appears to induce even rational agents to relinquish possibilities to improve their living standards, and thus stay mired in poverty.

Although these findings signal the importance of hope in the economics of poverty, several important caveats exist. First, psychological constructs like hope are not directly observable (i.e., latent) and do not readily lend themselves to empirical analysis. It is difficult for economists to conduct experiments, create theories, or enhance interventions without the ability to appropriately measure the intended constructs. For example, Banerjee et al. (2015) present results from six randomized control trials

of an integrated approach to improve the livelihoods among the very poor. In addition to receiving a productive asset, the participants also received services such as consumption support, training, and coaching. The approach resulted in an increase in self-employment income, as well as an improvement in the psychological well-being of the participants. The authors admit that the improvements in psychological well-being cannot be fully grasped by the existing economic theories and a more detailed psychological measurement is necessary to understand the underlying mechanism. Second, a systematic review on measuring hope shows that most studies were conducted in developed countries (Redlich-Amirav et al. 2018), so the measurement tools used do not necessarily translate well to the context of developing countries. The third caveat is that two-thirds of the studies on the measurement of hope were conducted in educational settings or in health facilities. And since these are very specific samples, the results may not be generalizable (Redlich-Amirav et al. 2018). Our data is constrained to rural Tanzania but covers a range of respondents aged 10-49 years. Fourth, measuring hope can be challenging because of measurement invariance, which is an underlying assumption in psychometric measurement tools that the items under consideration are understood in the same way by different populations. Understanding individual and group differences in measuring hope levels may be critical given that those differences have an apparent influence on the poor's economic behavior and outcomes. Currently, the issue of measurement invariance in the quantitative measurement of hope is unexplored and requires investigation (Bloem et al. 2018).

Given these caveats in the measurement of hope, using psychometric measurement tools that provide unbiased evaluation of psychological characteristics is critical to adequately assessing, planning interventions, and tracking development activities in the region. Development agencies seem to be recognizing this fact as they pay increasingly more attention to hope and a range of other latent variables.

1.0.1 Objectives

Taken together, the overall goal of this work is to contribute to the emerging literature on hope by addressing these caveats via two objectives. To do so we ask the following questions:

1. Are the survey questions measuring hope adequately?

2. Are different subgroups (e.g., men and women) perceiving the questions the same way?

To meet our objectives, our study team carried out a hope survey with more than 5200 individuals as part of a comprehensive household survey collected from 2816 households living in 98 villages in the Kigoma and Iringa regions, Tanzania. This is the preliminary stage of future field experiments designed to measure the economic and well-being effects of interventions aimed at changing people's intrinsic motivation in the region. The hope survey we use was adapted from the Scioli Hope Scale which defines hope as "a future-directed, four-channel emotion network comprising four constituent channels of mastery, attachment, survival, and spiritual systems" (Scioli et al. 2011). We use a short version of the survey consisting of 12 out of 40 items in the original Scioli Hope Scale, of which 8 are positively worded and 4 are negatively worded. The questions were carefully translated into Kiswahili. Respondents were asked to rate the extent of agreement with these items on a 4-point Likert-type scale with the options to strongly disagree, disagree, agree, and strongly agree.

We answer the first question in Chapter 5 where we apply the Graded Response Model (GRM) of Item Response Theory (IRT). Our aim in this chapter is to see whether questions in our survey measure hope as they are supposed to. We answer the second question in Chapter 5.3 where we use differential item functioning (DIF) analysis, which helps to meaningfully compare different sub-groups in terms of their tendencies to respond in certain ways to the survey items. Our goal in this chapter

is to determine how well the hope questions consider discrepancies between people of different demographic characteristics and their access to public services. [1]

To address our objectives, we use the baseline data that was collected for the Empowered World View (EWV) study by the World Vision (WV) team and its collaborators, Ifakara Health Institute, and the University of Alberta in 2020. The EWV is a behavioral change curriculum that seeks to change mindsets and improve empowerment and well-being in economically poor households and communities. The EWV aims to help economically vulnerable people and communities move away from poverty and dependency mindsets and towards empowerment and personal responsibility (*World Vision. Watoto Wetu 2018: World Vision Annual Report,* 2018).

The remainder of the thesis is organized as follows. Chapter 2 provides a background of the hope construct, a review of the theoretical and empirical literature on hope and its role in changing behavior, as well as a socio-economic overview of Tanzania, World Vision's work in the region, and brief information about the EWV program. Chapter 3 provides a conceptual framework followed by Chapter 4 where we present the overview of the study area and basic summary statistics on hope and demographics of the population. Chapter 5 presents the results of IRT and DIF. Finally, Chapter C concludes our findings with a discussion. Appendix A presents the results of the same analysis for self-efficacy and social cohesion questions and Appendix B presents a preliminary analysis of the multidimensionality of hope.

[1] Additionally, our IRT and DIF analysis for two other constructs - self-efficacy and social cohesion, which are other psychological constructs that play important roles in human behavior can be found in appendix C

Chapter 2

Background and Literature Review

The objective of this chapter is to situate our study in the existing economic literature of hope. To do so, we first review the empirical work that analyzed hope and how it might change economic behavior. We then review the hope literature focusing on developing countries, as well as discuss the distribution of hope across different groups of people. In the background section we discuss different definitions of hope, review the literature on how latent traits like hope have been measured, and provide a brief overview of Tanzania and the World Vision Tanzania's work. We use "latent trait" and "latent construct" interchangeably hereafter.

2.1 Literature Review

2.1.1 The Economics of Hope

Development economics has traditionally viewed poverty as a product of *external* constraints such as a lack of access to finance, education, healthcare, infrastructure, and technology (Lybbert et al. 2016). Thus, the solutions to poverty were also traditionally geared towards addressing external constraints. However, a greater recognition of the role of relieving one's internal constraints, such as an individual's hope, in alleviating poverty, gave rise to a new strand of development economic literature aiming to design new solutions to poverty based on the understanding of the role of hope and aspirations in individuals' behavior.

Economic research related to hope and aspirations has its origin in the work of anthropologist Appadurai (2004). Appadurai states that future orientation, an important aspect of hope, is a fundamental and hidden determinant of behavior that has radical implications for poverty alleviation and development. He views hope as a "navigational capacity", which is like a map that shows how to get to where you want to go in the future (Appadurai 2004). Appadurai's work laid the foundation for other economic models trying to understand the link between hope and economic development.

Following Appadurai (2004), Ray (2006) analyzed how poverty and aspiration or hope failures may be mutually reinforcing and built models that have been critical in the study of aspirations. Individuals compare themselves to their peers as well as those within their grasp (Ray 2006). Aspirations are shaped by the lives and accomplishments of "similar" or "attainable" individuals, or in other words, of a "local society". The "aspiration gap"—the difference between an individual's desired level of life and the standard of living he or she now enjoys—has an effect on their future-oriented behavior. Individuals with ambitions that are too close to their current quality of life, or with an excessively small aspiration gap, have little incentive to improve their welfare. On the other hand, people with goals that are too far removed from their current quality of life - those with excessive expectations – also lack motivation, as achieving an aspiration requires a great deal of effort, and so they may give up before they even start. Ray (2006) asserts that while an economically or socially polarized society may exhibit both of these characteristics, failure to achieve ambitions is distinct: a narrow aspirations gap results in the acceptance of one's life circumstances as a given (fatalism), whereas a wide aspirations gap results in dissatisfaction. Between these two extremes, an optimal aspiration gap exists that maximizes effort toward future-oriented behavior.

A related work by Snyder (2002) maintains that a key part of hope is that someone can see a clear link between what they do now and what they will do in the future.

In the absence of this element, even if the person has relatively low aspirations and a lot of agency to accomplish it, he or she may not always behave in a future-oriented manner. Furthermore, true agency and pathways may not always be the same as perceived agency and pathways (Snyder et al. 1991; Snyder 2002).

Following Appadurai (2004) and Ray (2006), and concepts of hope in psychology, development economists Lybbert et al. (2016) developed an economic model of hope. The framework is based on a reference-dependent utility that incorporates aspirations, agency, and pathways. It demonstrates how hope and ambition can influence internal development limits and potential poverty traps. They examine and differentiate "aspirational hope" from "wishful hope," arguing that whereas aspirational hope – "hope to..." is associated with growth and productivity, wishful hope – "hope that..." is an optimistic form of hope that results in disappointment and risky behavior. This conceptual framework is valuable in quantifying hope in poorer nations since the poor may lack agency and rely on outside assistance. Nevertheless, studies on the role of hope in poverty alleviation are mostly theoretical and lack empirical support (Duflo 2012). Lybbert et al. (2016) also talk about how low self-efficacy and internalized constraints might make people act a certain way. For example, a young girl thinks that jobs as engineers are not accessible to women, so she puts less effort into school. This internalization of constraints on pathways is different, however, from a scenario where low self-efficacy makes the girl think she cannot keep up with the grades she needs to get a degree. This could be a poverty trap since the effort required to determine what might be true limits lies beyond the equilibrium path. Sen (2014), also talks about internal constraints as a possible reason why a person might not be able to do what they want. People think the constraints on their way out of poverty are more restrictive than they really are.

Janzen et al. (2017) using theories of Appadurai (2004), Ray (2006), and Genicot and Ray (2017) discover that in rural Nepal, income and education aspirations are indeed related to the amount of income and educational achievement of people in an

individual's social circle. This discovery has implications for a potential mechanism for "spillover" effects. If the neighbor's livelihood gets better, then others may perceive that their livelihood improves soon as well. A deeper knowledge of the social dynamics that shape aspirations and impact behavior can help inform the design of policies and interventions to change the mindset of the people. Also, impact assessments of such initiatives should consider spillover effects, to be able to assess their full societal impact. Later in this work, we explore the validity of the items related to social cohesion that are also present in our survey.

2.1.2 Hope and Behaviour in developing countries

To understand what may motivate economists to study hope more closely, it is useful to review the literature on hope in developing countries. This literature focuses on understanding how hope affects the decisions of individuals and households and the outcomes of those decisions.

The role of hope can be useful especially in developing countries because of conditions created by a variety of conflicts, disasters, and social prejudices. Previous work shows that hope is a powerful motivator of human behavior in such contexts (Kleist and Jansen 2016a; Luthans and Jensen 2002; Snyder et al. 2000; Webb 2007). It is associated with less stress, higher life satisfaction, well-being, and improved quality of life. It is positively correlated with resilience, and self-efficacy, and is negatively correlated with depression and anxiety (Hutz et al. 2014; Rustøen, Cooper, and Miaskowski 2011; Snyder 2002). Becoming a non-smoker, eating more fruits and vegetables, and engaging in regular physical activity are examples of health-enhancing actions that have been correlated with hope (Berg et al. 2011; Nothwehr, Clark, and Perkins 2013; Anderson and Feldman 2020). Hope has been linked to mental health issues, such as anxiety, depression, suicide, substance abuse, stress, post-traumatic stress, and well-being (Gallagher and Lopez 2009), as well as academic achievement (Snyder 2002), productivity, and leadership in the workplace (Adams and Pulvers

2002).

When it comes to empirical research of hope and aspirations in the development economics, they were tested primarily in the domain of randomized control trials (RCTs). The RCT is a type of study in which people are randomly assigned to one of two groups: the experimental group, which gets the treatment that is being tested, and the control group, which gets an alternative treatment that is not being tested. The two groups are then followed up to discover if their outcomes differ. The trial's results and analysis are used to assess the intervention's effectiveness, or how much a treatment, procedure, or service benefits patients. RCTs are the most rigorous method for establishing a causal relationship between the intervention and the outcome (Sibbald and Roland 1994). Beaman et al. (2012) investigated whether observing female political leaders raises the aspirations of girls and their parents by using a self-reported set of five questions on educational attainment, future occupation, age of marriage, and leadership potential. By exploiting RCTs in West Bengal they found that the gender gap in parents' and adolescents' aspirations decreases due to a role model effect in villages where women were assigned to leadership positions. They say that aggregating aspirational questions in one index may not be legitimate because, as the questions may belong to different spheres of aspirations, they may confound each other's effects. Thus, after some validation of the survey questions, they excluded one question with the lowest correlation from the aspirations index (Beaman et al. 2012). An earlier study conducted by Jensen and Oster (2009) using individual panel data found that exposure to cable television affects the aspirations of women by lowering the tolerance towards domestic violence, higher school enrollment of children, increased women's autonomy, and decreased fertility among adult women. This happened not because television caused building a school or reduced school fees, but because television exposed rural households to urban lifestyles, values, and behaviors that are very different from their own (Jensen and Oster 2009). Adding to their study that such interventions can be effective in shifting perceptions and behaviors,

Tanguy et al. (2014) carried out a field experiment in Ethiopia where poor people listened to stories of how people similar to them escaped the poverty. Six months later, they found that the treatment group had experienced significant positive impacts on aspirations index, actual savings, investment in education, and time allocation for business. They argue that the intervention is not linked to updating beliefs because of new information, but rather a "vicarious experience" of how someone similar to them could get through poverty due to hard work and persistence and this inspires them to have more hope in their children (Tanguy et al. 2014). These results support the hypothesis set by Appadurai (2004) and Ray (2006) that aspirations, although an individual attribute, respond to collective influence. Despite the huge economic and psychological barriers that the very poor face, even relatively minor changes in their circumstances can considerably raise their chances of escaping poverty. For example, providing a single asset (a cow or a sewing machine) to very poor people in West Bengal resulted in a sustained increase in consumption that exceeded the asset value, and the beneficiaries increased their overall income in ways not entirely traceable to the initial assets (Duflo 2012). The transfer's effects on economic behavior and emotional well-being much surpassed what the researchers could have anticipated based on the transfer's economic value alone.

The economic literature increasingly recognizes that a person's positive and negative outlook about the future can explain economic behavior. For example, studies found that optimistic people are happier and more resilient to negative shocks (Ekici and Koydemir 2016). But there is some evidence that excessive optimism may be bad for health by leading to disappointment or social disillusionment (Easterlin 2001; Kleist and Jansen 2016b; Arampatzi et al. 2018). Also, a strong correlation has been discovered between pessimism and negative outcomes like early mortality and low labor force participation (O'Connor and Graham 2018; Graham and Pinto 2019). In behavioral economics, a growing body of studies examines the relationship between self-esteem and economic outcomes. For example, Bénabou and Tirole (2003) demon-

strate that empowering and encouraging an individual can increase self-esteem, which can result in increased achievement (Bénabou et al. n.d.). Darolia and Wydick (2011) discover that measures such as parental praise aimed at increasing self-esteem result in academic accomplishment above what inherent ability would indicate for university students and the more self-esteem a child has, the more likely he or she is to try harder, believing that effort put into a task will pay off in the long run. Students who received gifts of large spending allowances and cars showed lower levels of effort and academic achievement (Darolia and Wydick 2011). Other related work by Duggal et al. (2016) associates hope with a positive mood that has direct impact on human health (Duggal, Sacks-Zimmerman, and Liberta 2016). Genc and Arslan (2021) link hope to subjective well-being, a concept that is gaining increasing recognition in economic studies. According to the findings of a study conducted on college students in Turkey, young adults who experience high levels of stress because of the coronavirus have less optimism and hope, resulting in lower subjective well-being (Genç and Arslan 2021). Marlon et al. (2019) claim that both hope and doubt are strong determinants of climate change mobilization via political behavior and support for greenhouse gas mitigation policies. They argue that hope is not always good, and doubt is not always bad; the mix of constructive hope and doubt can be motivating, whereas false hope (e.g., wishful thinking) and fatalistic doubt (e.g., beliefs that there is nothing humans can do) can lead to avoidance, distancing, and inaction (Marlon et al. 2019).

2.1.3 Is hope the same for all?

Hope is not evenly distributed among different groups. It is affected by a person's position in society, exposure to other practices, and ways of life, access to and control over resources, and, most crucially, knowledge. For example, some women in Bangladesh aspire to eat two meals a day, look clean, and wear good clothes. They do not use loans available, because they do not have the capacity to start a business

or are willing to take risks. But they have ambitious aspirations for their children, they want them to be educated. On the other hand, the wives of pond owners and other market operators have high aspirations not only for their children, but for themselves as well. They want to accumulate and run their businesses on a bigger scale, take ponds on a lease, use capital to hire other people, build a cement-brick house, buy a TV, and so on (Nathan 2005). Aspirations among women and men can also differ due to gender division of responsibilities, their economic roles in the household, and cultural practices. Another example, the poor don't have the opportunities or routes they need to reach their goals, because they have lower chances of seeing and practicing that navigational capacity. Their aspirations are drawn from their own experience and are strongly affected by their social environment. The poor can wish, hope, or aspire, but the pathways (ways of achieving goals) from their wants to intermediate real experiences are weaker. The poor will likely under-invest in the future, not because of their intellectual deficit but because of their low capacity to aspire. Thus, social climbing among the poor might be hindered by internal constraints such as low hope, aspirations, and self-efficacy (Appadurai, 2004).

Snyder (1994) developed a scale for adults' and children's hope to determine whether there are differences in hope scores by ethnic group, gender, or age. This aided in the development of interventions tailored to specific groups. For example, despite efforts to equalize samples based on socioeconomic background, the study in hope between public and Catholic schools discovered considerable differences in hope between schools. Separate examinations of the samples revealed apparent racial distinctions (McDermott et al., 2002). De Sá et al., (2019) examined gender and age variations in depression measurement among Brazilian college students using item response theory and differential item functioning tests and discovered that women and younger students are more likely to suffer from depression than older and males. However, results indicated that statements such as "crying" cause women to overestimate their symptoms and older students to indicate "loss of interest in sex". As

such, it may be critical that measures used to assess hope incorporate gender and age differences. For instance, males may be less likely than females to endorse spiritual views, which may distort the predicted degree of hope by gender, exaggerate the level of hope among females, or vice versa.

2.2 Background

2.2.1 Roots, definitions, and hope in other disciplines

We see that each scientific field takes a unique approach to the measurement of hope. For example, the most widely researched model of hope in the psychology literature is Snyder's (2002) Hope Theory. According to Snyder's definition, someone who has hope has a goal in mind, sees a way to get there, and thinks he/she has the agency (power) to get there (Snyder et al. 2000). However, Snyder's theory has been criticized for being overly individualistic (Du and King 2013), for focusing too much on personal control and action and too little on trust (Tennen and Tennen 2002), for failing to distinguish enough between hope and optimism (Aspinwall and Leaf 2002), for ignoring the significance of emotion and for departing significantly from how individuals perceive hope in daily life (Tong et al. 2010). As a result of these and other criticisms, a few more theories have been developed that focus on spiritual, social, open-ended hope and multidimensionality of hope (Scioli et al. 2011; Du and King 2013). One that has been used in many health studies comes from Herth(1992). She describes hope as "a multidimensional dynamic life-force characterized by a confident yet uncertain expectation of achieving good, which to the hopeful individual, is realistically attainable and personally significant" (1992). To measure hope among patients, the Herth Hope Index relies on expectations, a positive attitude toward the future, and the social context in which hope occurs. Later on, Schrank et al. (2011) came up with a way to measure hope that combines instruments from psychology (the Snyder Adult Trait Hope Scale) and health sciences (the Herth

Hope index and the Miller Hope Scale). Hope research in sociology, anthropology and political scientific fields underlines how hope is a part of cultural, societal, historical, and political structures, not something that is detached from the rest of society (Kleist and Jansen 2016a). Webb (2007) came up with a bigger, more interdisciplinary theory of hope that starts with the fact that there are many different and sometimes even conflicting descriptions of hope.

In this thesis, we use Scioli et al.'s (2011) definition of hope. Scioli et al. (2011) created an alternative model of hope as a future-directed network of emotions using literature from psychology, philosophy, theology, spirituality, and nursing. They explain hope in terms of a "network" of four interacting components such as mastery, attachment, survival, and spirituality. It is broad and integrative, going beyond goals to include social support and life meaning (Scioli et al. 2011). In general, hopeful individuals are attached, empowered, and skillful at controlling their responses to stress and loss (Scioli and Biller 2009). Given that hope can be understood in these different ways, and analyzed from different viewpoints, we emphasize our analysis of the Scioli et al (2011) conception of hope is one of several possibilities.

Before moving to the next chapter of the work, we will give a brief overview about Tanzania, discuss World Vision's work in the region, and look at the goals of the Empowered World View (EWV) training program.

2.2.2 A brief overview of Tanzania

Tanzania, East Africa's second-largest country, has one of the region's lowest population densities. It is a resource-rich country that has great potential in the agricultural and mining industries. Tanzania achieved a significant milestone in July 2020, when it formally transitioned from low-income to lower-middle-income country status. Tanzania's success is a result of sustained macroeconomic stability that has helped the country grow, as well as its rich natural resources and strategic geographic location (*World Bank,* 2021).

Kigoma is a region in northern Tanzania, located on the shores of Lake Tanganyika. It borders Burundi to the north and the Democratic Republic of Congo to the west. The Kigoma region has a long history of hosting refugees. The majority of the hosted refugees are from Burundi and the Democratic Republic of Congo who fled their countries in the 1990s and again in 2015-2017.

Iringa is bounded to the north by Singida and Dodoma, to the east by Morogoro, to the west by Mbeya, and to the south by Njombe. It is predominantly agricultural, and it has the country's second-highest per-capita GDP. This is reflected in higher literacy rates of people aged 15 and above, in Iringa it is 81.9 % while in Kigoma 76% (*National Bureau of Statistics* 2014). Farming is the most common occupation on average among the working population (63%) followed by elementary occupations and; service and shop sales workers (each employing 6%), craftsmen (5%), technicians and associate professionals; and other not specified (4%) each, street vendors (3%), livestock keepers and professionals (2%) each. In Kigoma farming as an employment status accounted for 78% while in Iringa 70.3%. The issue on ownership of certain assets builds poverty monitoring indicators. The percentage of households who own assets is about twice as high in Iringa as it is in Kigoma. (*National Bureau of Statistics* 2014)

Tanzanian's growth trajectory had many achievements as well as challenges. The country has a reputation as a stable and peaceful country with promising GDP growth and a thriving private sector that boosted living standards. Nonetheless, these accomplishments have not resulted in long-term economic success and political equality for all. According to the multidimensional poverty index from 2015/2016 data, which identifies multiple overlapping deprivations suffered by individuals in 3 dimensions: health, education, and standard of living, 55.4 % of the population in Tanzania was classified as multidimensionally poor. Tanzanians wanting to overcome poverty face an uphill battle due to a lack of employment opportunities, especially in rural areas. In 2019, 15% of Tanzanian youngsters were neither working nor enrolled in

education (*World Faiths Development Dialogue (WFDD). Faith and Development in Focus: Tanzania. Berkely Centre for Religion, Peace and World Affairs, Georgetown University,* 2018).

In Tanzania, undernutrition remains one of the country's biggest human development challenges. (IRIS 2017). According to estimates, 450,000 children in Tanzania are acutely malnourished or wasted, with 100,000 suffering from the most severe form of malnourishment. Nationally, chronic malnutrition affects 34.7% of under-fives, while 11.5% suffer from severe stunting. In Kigoma chronic malnutrition exceeds 40%, while in Iringa more than half the children are chronically malnourished (51.3%) (*IRIS,* 2017).

Malnutrition is also a great barrier to academic achievement and a major obstacle to economic progress because of the negative impact it has on people's health, capacity to learn, and overall productivity. Despite significant progress, malnutrition rates remain high, particularly among rural households (*Tanzania National Nutrition Survey,* 2018). In the mid-2000s, nearly all younger children in the country were enrolled in elementary school, which is a big step forward for the education system of Tanzania. Nonetheless, school attendance and enrollment rates in Tanzania have shown that there are bigger social discrepancies in the country. Children from poorer households are three times more likely to drop out of primary school than those from wealthier households. This discrepancy is exemplified geographically, with rural areas often having lower enrollment rates than urban areas (*UNICEF. Education Fact Sheet:Tanzania,* 2017). There are still children, especially those from poor households, who do not attend school for several reasons, such as long distances between school and home and the opportunity to earn money through menial labor instead (*UNICEF. Education Budget Brief,* 2020). In 2019, Iringa had greater enrolment rates than the national average, while Kigoma had lower enrolment rates. Kigoma also has a gender disparity that favors boys. (*The World Bank Group. Tanzania Gender Assessment* 2022). Also, school retention and quality of learning are ma-

jor problems in the educational systems. It is hard for children to learn because of overcrowding and a lack of teacher training (*UNICEF. Education Budget Brief,* 2018).

Currently, 1.6 million Tanzanians—or 4.6% of the population are living with HIV, 72,000 of whom were newly infected in 2018. Prevalence rates vary widely by region, HIV prevalence rate in the province of Iringa is up to 11.3% among adults aged 15 years and older, while in Kigoma is 3% (*Tanzania HIV Impact Survey 2016-2017: Final Report,* 2018). Women are more likely to get infected than men because they marry earlier, have older partners, and have more trouble negotiating safe sex. It should be noted that faith-based hospitals are used by about 40% of the population, even though three-fourths of hospitals are public (*World Faiths Development Dialogue (WFDD). Faith and Development in Focus: Tanzania. Berkely Centre for Religion, Peace and World Affairs, Georgetown University,* 2018).

In Tanzania, 31% of girls get married before they turn 18. This makes Tanzania the eleventh country in the world with the most child brides. Interesting that Iringa, one of our sites has the lowest percentage, about 8% of child brides in the country (*Tanzania, Girls Not Brides,* 2021).

2.2.3 World Vision in Tanzania

In Tanzania, World Vision's work is mostly focused on livelihoods and income generation. World Vision has been providing smallholder farmers with the skills, technology, and finance necessary to create income since the 1980s (*World Vision Tanzania,* 2022). World Vision developed over 1,300 savings groups with about 28,000 members of which 69% are women in 2018, provided loans to over 4,300 farmers through Vision Fund Tanzania, and taught over 20,000 people in efficient and sustainable agricultural practices, and over 160,000 people accessed a basic drinking water source in 2018 (*World Vision. Watoto Wetu 2018: World Vision Annual Report,* 2018).

Empowered World View training

The World Vision, through its EWV training, aims to make an impact both via changing beliefs and through shifting from relatively fixed and pessimistic mindsets to more optimistic growth mindsets. Thus, recent studies on mental health (e.g. Reis et al. (2019), religion (Scioli et al., 2011), education (eg. Das and Zajonc, 2010), development interventions (eg. Lybbert and Wydick, 2018; Duflo and Banerjee, 2011), self-efficacy (Schwarzer and Jerusalem, 1995) and gender analysis (Malapit et al., 2019) were used to construct a mindset section of the survey that the World Vision / IHI / University of Alberta research team is conducting to measure the success of the delivered development programs. The research team began by selecting a series of English-language questions that had been tested in a variety of settings, and modified them for the Tanzanian context to administer in the field using tablets.

The method of distributing EWV knowledge entails preparing community leaders to conduct EWV workshops and then assisting these facilitators to train people who are influential in their communities on EWV concepts. Involving community leaders in EWV implementation is critical to ensuring community acceptance of the EWV ideology of fostering individual efforts to improve one's conditions. EWV considers faith leaders, who are among the most influential in local communities, leaders from farmers' associations, who will inspire their peers to implement new ideas to increase crop production and income, and community leaders, including leaders of women's and youth groups and health extension workers to help identify creative ways to use their skills to shape their futures (*World Vision. Breaking the chains of poverty with Empowered World View,* 2017).

Chapter 3

Conceptual Framework

In this thesis, we aim to contribute to the emerging economic literature on measuring latent variables, or hypothetical constructs that cannot be measured directly. Specifically, we consider hope, using data from a survey conducted in two regions of Tanzania. We have two objectives, each focusing on two issues arising in the measurement of latent constructs. In this section, we will provide a summary of Item Response Theory (IRT), and its assumptions. Then we will introduce the specification of IRT called the Graded Response Model (GRM) which we will use to meet our first objective. Finally, we will introduce a method called Differential Response Functioning (DIF), which will be used to meet our second objective.

3.1 Psychometric measurement of latent constructs

A good measurement instrument of latent constructs includes a list of one or more questions or statements that collectively result in the accurate assessment of the latent trait. In psychometrics, these statements or questions are referred to as "items". In psychometric testing, item construction is an essential component (Furr 2014). Items should be worded in a way that does not lead respondents to answer in a certain way. Items should also refer to a single concept or issue to ensure clarity. Another important feature of a good item is the ability to accurately discriminate respondents based on their level of the latent trait. For instance, in diagnosing depression, a ques-

tion such as, "Do you feel constantly blue?" does not provide an accurate assessment of the condition. Even if the answer "yes" is likely to identify most individuals who are really depressed, it is also likely to include those who are experiencing a difficult time due to a recent loss and do not meet the criteria for clinical depressive disorder. Also, depression is a complex illness, which is more than feeling sad. We may be overlooking essential features of depression if we merely ask one question and fail to recognize those who are depressed as a result (*Core Guide: Measure Development,* 2020).

It may be interesting to review the process of developing measurement tools for latent constructs typically used in the psychometric literature. It usually starts with defining the latent construct, reading relevant literature, and learning about the population through intensive in-depth interviews. Then, if the existing scale was previously verified, the researcher utilizes the existing scale or modifies it, or designs/combines their own scale to measure the construct based on the others' work. The items adapted from an existing scale are translated and back-translated if they are not in the language of interest (Brislin 1970). For instance, a back-and-forth translation of the word "feeling blue" in the measurement of depression would guarantee that the translated item makes use of a culturally and linguistically related phrase to help express the same meaning. Thus, in contrast to development economics which often uses the same measurement tools across different contexts, a best practice in psychometrics is to adapt existing measurement tools to new cultural settings, contexts, and different populations prior to analysis.

A successful measurement tool of a latent construct requires statistical procedures such as validity and reliability. It is important to ask, "is this construct measuring what we intended?", which is a validity question. Methods such as principal component analysis, factor analysis, and item response theory are used to look at participants' responses and figure out the underlying dimensionality structure, or the latent model. Reliability asks "is the measurement consistent?". A visual representa-

tion of these principles is provided in the figure 3.1 below. To test internal reliability methods such as Cronbach's alpha, the omega coefficient, and split-half are used.

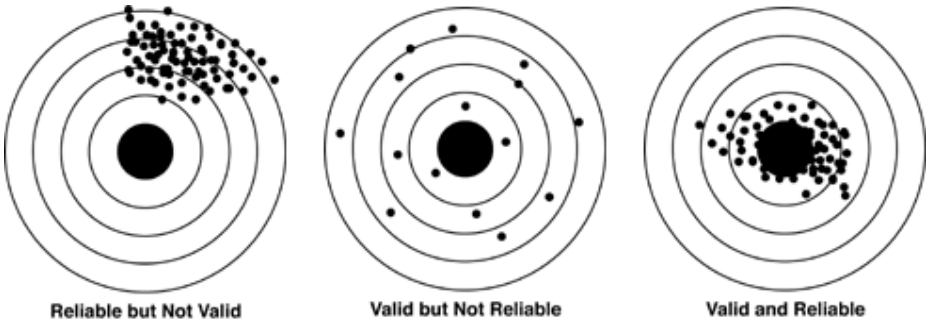


Figure 3.1: Visual depiction of validity and reliability (CCNMTL 2003)

3.2 Item Response Theory

The method we are using to meet the first objective of the thesis is called Item Response Theory (IRT). IRT is a probabilistic model of how a participant would respond to any given item(s) depending on their degree of the latent trait. In our case, an individual's hope level (latent trait denoted by θ) is the probability of endorsing the item. So, the more an individual's hope, the greater the probability of providing an answer "strongly agree" to positive items and "strongly disagree" to negative items. The probability of endorsing them increases monotonically as the respondent's hope level increases (Embretson 2000). The two-parameter logistic (2PL) model is a frequently used IRT model for dichotomous items (Birnbaum, A. 1968). This function depends on two parameters - the item's discrimination (a) and difficulty (b). Difficulty is the minimum latent trait level required to correctly answer an item. It is also known as "location" for their location on the difficulty range. Discrimination (slope) is an item's ability to distinguish between individuals with low and high latent trait levels. We may build an item characteristic curve (ICC) based on the item's difficulty and discrimination to determine a respondent's chance of endorsing the item in relation to his or her latent trait level.

In IRT, the latent trait and item difficulties are typically rated on a standardized

metric across a sample of respondents, with means of 0 and standard deviations of 1. For example, a person with a hope level of 0 has an average level of hope, and a person that has a hope level equal to 1 is 1 standard deviation above the average. An item with a difficulty level equal to 0 has an average difficulty, while an item with a difficulty level of 1 is a relatively difficult item. If an item has a difficulty of 0, then a person with an average level of hope will have a 50% chance of agreeing with the item. If the discrimination is positive and high, then the item is consistent with the trait we are measuring. Low levels of discrimination indicate that not only hope, in our case, is needed to positively endorse the item, but also something else. If the discrimination is 0, then it is unrelated to the trait level. And if it is negative, then it is inversely related to hope, meaning that high levels of hope make respondents less likely to agree with the items. A perfect item would be highly discriminating and have threshold settings that span a wide range of the latent variable (Embretson 2000).

Ideally, items should provide strong information across the whole spectrum of the latent trait. In the worst case, all our items give little information for any of the hope levels. We also want to avoid a situation when our items' informativeness is concentrated around the same levels of hope. When the IRT model provides information functions for each item, it is called an item information function (IIF), which shows the reliability of an item at different points of the trait. Items need to capture a range of hope levels (low, medium, or high). Items may differ in two respects. For example, a very easy math question will provide little information on who has high ability since most will answer it correctly; similarly, a very difficult question will carry little information about which individuals have the low ability because it is too difficult for them.

A graph of information functions for all items taken together is called a test information function (TIF). TIF provides the reliability of all items together on the same graph. IRT additionally offers the TIF's standard error which is the inverse of

the TIF; therefore, both are measures of the test's precision at various trait levels. IRT shows the standard error of the TIF, which indicates the precision of the test at different trait levels (Embretson 2000).

3.2.1 IRT Assumptions

The proper use of IRT requires a set of assumptions to hold (Embretson 2000; Nguyen et al. 2014). The first assumption is monotonicity which means that the probability of a person endorsing an item increases as the person's latent trait level increases. The second IRT assumption is uni-dimensionality, which refers to the notion that the items on a scale all measure the same thing: that they are all impacted by a single latent feature. The third IRT assumption is item invariance. If the item invariance assumption is satisfied, respondents with similar latent traits will get similar scores on the same or a completely different test designed to measure the same trait. In addition, respondents who have different latent traits will keep the same difference in their scores if they take the same test. The fourth and final IRT assumption is local independence. This hypothesis refers to the responses of subjects with the same ability or trait level on different test items being independent of each other. They are uncorrelated after controlling for the latent trait. Local independence and uni-dimensionality are different concepts, but they imply each other. If local independence is satisfied, the probability of the correct response to the same test item is the same for all subjects of the same level, meaning the test is unidimensional (Reese 1999). Therefore, satisfying one assumption implies satisfying the other as well.

3.3 Graded Response Model

An important step in IRT analysis is choosing an appropriate measurement model. Polytomous IRT models are required to reflect the nonlinear relationship between a respondent's latent trait and the likelihood of responding in each category for multiple-category item-response data. When item responses are coded into more

than two ordered categories, Samejima's graded response model (GRM) is appropriate (Samejima 1969). The GRM is specified as follows:

$$P(X_{ri} >= k|a_i, b_{ik}, \theta_r) = 1/(1 + \exp(-a_i(\theta_r - b_{ik})) \qquad N(0,1) \tag{3.1}$$

where X represents respondents r's response to each item i. The GRM models the probability of a respondent selecting a score at or above each item score category. With a four-category Likert scaled item scored from 1 to 4, the probability of scoring at or above 1 is 1, but for score points $k = 2 \ldots 4$, the probability is modeled as four logistic functions rising from 0 to 1 across the latent score scale, theta. Here, θ_r represents the latent score for each respondent r (hypothesized to indicate a respondent's level of hope), a_i is the information parameter for each item i, and b_{ik} represents the location parameter for each item i, and score category k. The information parameter (a_i) indicates how well an item can distinguish between respondents with very similar hope levels. The location parameter (b_{ik}) indicates whether respondents need a higher or lower level of hope, θ_r, to respond at or above that level k. The θ_r parameters can be interpreted on a standard normal scale, where –1 and +1 are one standard deviation below and above the mean, correspondingly. In a strict sense, a_i is the increase in the log of the odds of scoring at or above item i's categories k, for each 1-standard-deviation-unit increase in θ_r. The location parameters, b_{ik}, indicate where a respondent with $\theta_r = b_{ik}$ has a 50% chance of scoring at or above category k of item i. These parameters are invariant between item and respondent populations.

3.3.1 Model choice and model fit

When designing our IRT study, we can assume that all the items are equally effective at discriminating between respondents, or we can assume that each item has a unique discrimination parameter. Thus, we are comparing a "constrained" model to an "unconstrained" model - and, as we will see, the difference between these models in terms of their fit to the data can be explicitly tested. To compare the model fit

statistics of constrained and unconstrained models, we use an ANOVA (Analysis of Variance) that evaluates the difference between the two models.

3.4 Differential Item Functioning

The second objective of this thesis is to find out if the hope questions are understood the same way by individuals of different backgrounds based on gender, age, region, religion, food security, access to water, and treatment groups. For this purpose, we will use Differential Item Functioning (DIF) analysis, which will help to meaningfully compare different sub-groups. Following the estimation of IRT item attributes, we can see if the items' properties (e.g., item discrimination and difficulty) are invariant across groups. In item analysis, IRT is a helpful analytical technique for detecting group differential item functioning (DIF) (Thissen 1993). Researchers use DIF analysis to investigate the potential that an item may operate differently for one focal group (e.g., females) than for another reference group (e.g., males) despite adjusting for variations in the assessed construct between the two groups. Typically, the reference group refers to a group presumed to be in an advantageous position, while the focal group is suspected to be at a disadvantage (Camilli 2006). When an item has a DIF, people with the same characteristic level but belonging to different groups (e.g., males versus females) have a varied chance of responding to the item. DIF items are sometimes referred to as "biased items" since they lead one group to score higher on a measure than the other merely due to their group membership.

DIF can happen for several reasons, including different interpretations, group norms on how to respond (e.g., avoiding extreme categories, such as "strongly agree" or "strongly disagree"), or other things such as how items are administered. As a result, assessing for DIF can be very useful in finding items that show cultural bias, which can then be changed or removed. Finding DIF is generally undesirable since such items might risk the validity of an instrument used to assess the trait levels of individuals from various populations or groups. On the other hand, DIF may lead to

a greater understanding of the psychological differences that underpin subgroup responses, as well as a future study into mechanisms for such differences. The presence of a DIF implies a conditional difference in item performance across groups, but it does not reveal the source of the DIF. Understanding the cause of DIF is important because it helps test developers better understand item and group properties that are responsible for differential test performance, and thus whether the DIF represents a fair or unfair advantage on an item. For example, a depression index has a question about the frequency of crying (De Sá et al. 2019). However, if women endorse this item more than men, keeping the level of depression constant, there is DIF. Therefore, the depression index will be biased towards diagnosing women more frequently than men if the item containing a DIF is not addressed. We will conduct a DIF analysis of our items as the second objective of the thesis.

In addition to considering DIF for focal and reference groups, we also look at the DIF between villages comprising the "control" and "intervention" groups which were pre-selected for the future phases of the study involving experiments. The control group consists of 7214 individuals living in 50 villages, and the intervention group consists of 7038 individuals living in 48 villages.

Chapter 4

Context and Data

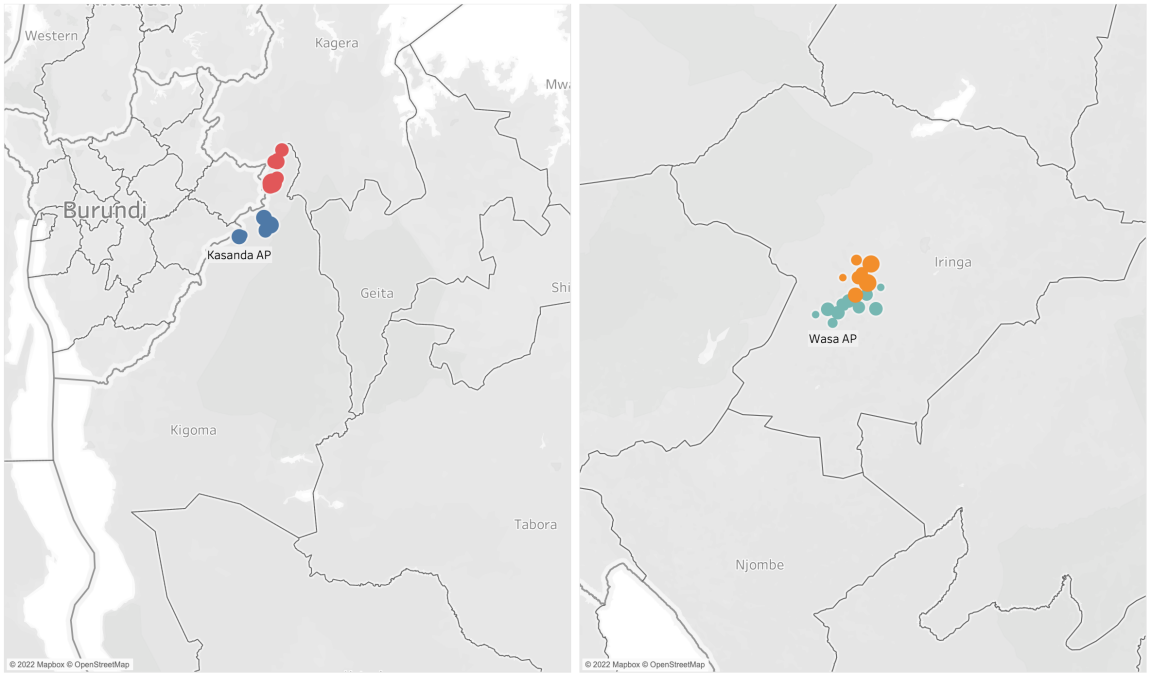
In the previous chapter we presented a conceptual framework of our study. In this chapter we discuss the context of the study and provide an overview of the data we use.

4.1 Overview of the study area

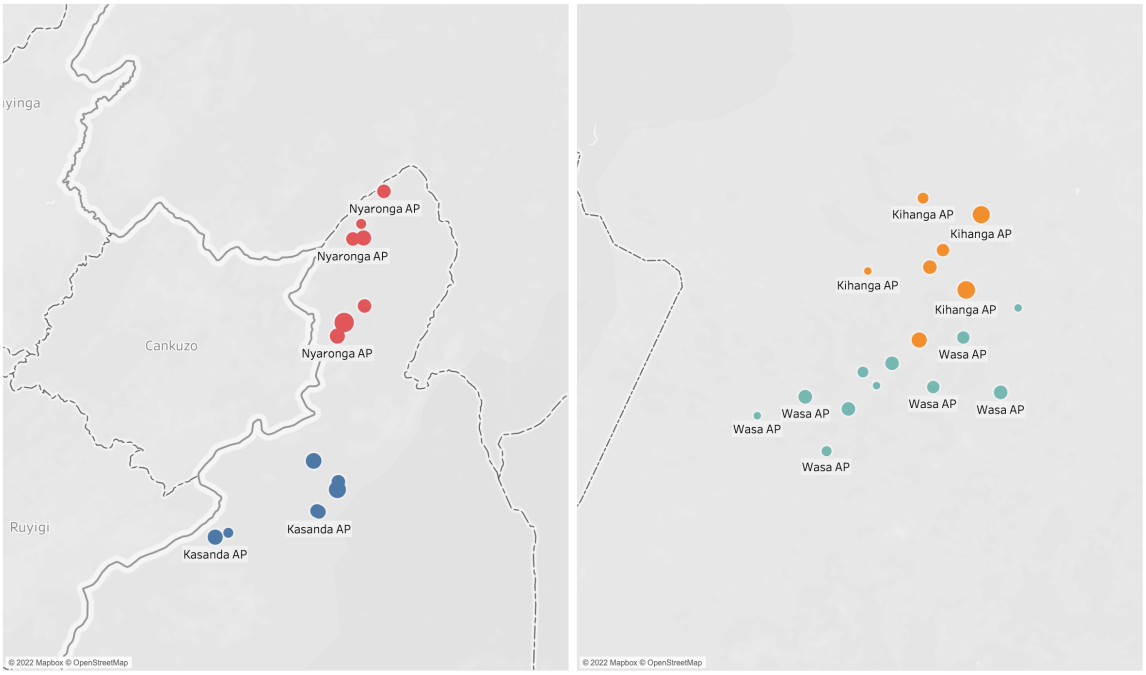
This study evaluates the baseline data that was collected for the mixed-methods study of Empowered World View (EWV) by the World Vision (WV) team and its collaborators, Ifakara Health Institute and the University of Alberta. Empowered Worldview (EWV) is a behavioral change curriculum that is based on biblical principles. It is meant to change mindsets and improve empowerment and well-being in economically poor households and communities by changing people's mindsets and changing their behavior. The EWV intervention is meant to help economically vulnerable people and communities move away from poverty and dependency mindsets and towards empowerment and personal responsibility.

The large survey data collected in July-August 2020 and an endline data collection scheduled for July-August 2022 are aimed at identifying the effectiveness of the EWV model and the types of improvements that EWV contributes to communities, as well as the paths to those changes. EWV will be implemented in four World Vision Area Programmes (APs) namely Kihanga, Wasa, Kasanda, and Nyaronga, located in

two regions in Tanzania: Iringa in the Southern Highlands Zone and Kigoma in the Western Zone. Figure 4.1 shows the area of study coverage.



(a) fig 1



(b) fig 2

Figure 4.1: This map is a depiction of villages located in the study areas in Kigoma and Iringa. Kigoma: Kasanda AP (Control) and Nyaronga AP (Intervention; Iringa: Wasa AP (Control) and Kihanga AP (Intervention)

The baseline survey is divided into four sections. First is the household data,

which consists of questions on household demographics, characteristics of all household members, household poverty and economic resilience, food supply, water and sanitation, and disaster management. In the second part of the survey, questions were asked of adolescents aged 10 to 19 years about their rights and protection, well-being, spiritual beliefs, and relationships with others. The third section consists of mindset questions asked from respondents aged 10 to 49 years. The choice of 49 years as the upper bound was made because that is the highest age of an adult with children under the age of 5 years. The fourth section is asked of all women about pregnancy and childbirth. Mothers or caregivers of children under 5 had additional questions about their children's health. Survey data consists of close-ended responses, either continuous or categorical.

A two-stage cluster random sampling approach was used, with the first stage consisting of village selection and the second stage consisting of household sampling within each village. Villages located at the border of the control and treatment areas were removed to reduce spill-over effects between treatment and control. A total of 98 villages were randomly selected, and a total of 2814 households were interviewed from those villages.

In Iringa and Kigoma, 1414 and 1400 households were interviewed. Our goal was to interview all household members aged greater than 10 years. The average household size in Iringa is 4.6 people, whereas in Kigoma it is 5.3 people. 99% of households in Iringa are Christian, while in Kigoma 88% are Christian and around 5% are Muslim. The breakdown of respondents by age and gender showed that in the category of adolescents (10-19 years old) males and females are almost equally represented, whereas in higher categories (20-49 years old) we have significantly more females than males. Half of the males are concentrated in the adolescent category. The females are more evenly spread out across age groups. In making inferences about males, we bear in mind that almost half of them are 19 or younger.

Gender	Female	3130 (59.9%)
	Male	2094 (40%)
Age (10-49 years)	Mean	26.09 (SD 11.76)
Region	Iringa	2756 (52.8%)
	Kigoma	2468 (47.2%)
Education	Total numbers of adolescents responded	491
	Never attended	20.60%
	Primary	55.78%
	Secondary	13.44%
Main Source of income	Sale/exchange of own produce (farming)	74.69%
	Wage employment (working for someone else)	11.33%
	Petty business/Vending/Trading	5.44%
	Labour (self-employed/e.g. driver)	3.24%
	Direct selling (e.g. cosmetics)	2.47%
	Other	3%
Marital status	Married	52.50%
	Not married	44%
	Divorced	3%
Religion	Catholic	62.75%
	Protestant/other Christian	29.90%
	Muslim	2.60%
	No religion	3.64%
	Other	1.11%
Total number of respondents		5,224

Table 4.1: Summary statistics concerning the socio-demographic characteristics of the respondents

Summary statistics of our sample in Table 4.1 indicate that 60% of respondents to the mindset questions are females, while 40% are males. The average age of respondents in our sample is 26. Only 491 adolescents between the age of 10-19 years answered the question about the highest educational level they have, and 20.6% of them said that they have never attended a school. 74.7% of the households dedicate themselves to the sale/exchange of their own produced products from farming.

11.3% of a household's main income is from wage employment. Non-agricultural self-employment is relatively rare; around 11% of households have income from self-employed labor, direct selling, and petty businesses.

The Scioli's Hope Scale in Table 4.2 is a 12-item self-reported measure of hope, with 3 items in each measuring four components viz., mastery (e.g "making progress towards important goals"), attachment (e.g., "feel loved by someone"), survival (e.g., "can handle difficulties"), and spirituality (e.g., "my spiritual beliefs have empowered me"). In each component, there is a negative item like "worry that someone may betray me". Respondents are asked to rate the extent of agreement with these items on a 4-point Likert-type scale ranging from strongly disagree to strongly agree. The order in which respondents were asked these 12 hope items and frequency distribution of responses are presented in the table 4.2.

	Hope Items	Strongly disagree	Disagree	Agree	Strongly agree
1	I worry that someone may betray me.	45	180	2680	2477
2	There are people in my life that I completely trust.	79	302	2868	2134
3	I worry that someone may betray me.	894	1971	1803	696
4	I'm running out of options for improving my life.	514	1571	2309	893
5	I'm making progress towards important goals	51	247	3246	1790
6	I have a purpose in life.	63	520	3242	1524
7	I can handle any current or future difficulties	337	725	3165	1087
8	The future will bring opportunities for a better life.	115	657	3455	1100
9	I have doubts about achieving those things that really matter to me.	494	1688	2439	713
10	My faith in a higher power gives me the strength to pursue my dreams.	84	324	3114	1818
11	My spiritual beliefs have empowered me to succeed in life.	87	255	2979	2022
12	I have never felt close to any kind of spiritual force or presence.	1289	2289	1425	329

Table 4.2: Frequency distribution for hope questions

Table 4.3 contains descriptive statistics for each of the hope items. The means for positive items fall into the categories of "agree" and "strongly agree" except for Item 5. The means of the negative items fall in between "agree" and "disagree".

Positive Items					
Note: 1-Strongly disagree; 2-Disagree; 3-Agree; 4-Strongly agree	Mean	SD	Skew	Kurtosis	Correlation with total score
1. I feel loved by someone (attachment)	3.41	0.6	-0.72	0.79	0.55
2. There are people in my life that I completely trust (attachment)	3.31	0.65	-0.74	0.95	0.52
3.I'm making progress towards important goals (mastery)	3.27	0.59	-0.43	0.86	0.55
4. I have a purpose in life (mastery)	3.17	0.63	-0.42	0.52	0.58
5.I can handle any current or future difficulties (survival)	2.95	0.77	-0.74	0.63	0.34
6. The future will bring opportunities for a better life (survival)	3.04	0.65	-0.52	1.01	0.57
7. My faith in a higher power gives me the strength to pursue my dreams (spirituality)	3.25	0.64	-0.64	1.05	0.6
8. My spiritual beliefs have empowered me to succeed in life (spirituality)	3.3	0.64	-0.74	1.23	0.58
Negative Items					
Note: 1-Strongly agree; 2-Agree; 3-Disagree; 4-Strongly disagree					
9. I have never felt close to any kind of spiritual force or presence (spirituality)	2.85	0.86	-0.29	-0.64	0.34
10. I have doubts about achieving those things that really matter to me (mastery)	2.36	0.83	0.21	-0.48	0.37
11. I'm running out of options for improving my life (survival)	2.32	0.87	0.23	-0.6	0.32
12. I worry that someone may betray me (attachment)	2.57	0.92	-0.05	-0.83	0.36
N=5229					

Table 4.3: Item Mean, Standard Deviation (SD), Skewness, Kurtosis for hope items

Most of the items are skewed left, meaning that the participants tended to agree with the positive hope items and disagree with the negative hope items, which indicates high levels of hope overall. It is important to note that the negative items, Items

9-12, are reverse coded, meaning that for these items a score of 1 means "Strongly agree" and a score of 4 means "Strongly disagree". By observation, the positive items have a higher mean than the negative items despite the reverse coding. This may already suggest that the positive and negative items may address slightly different constructs. Furthermore, the last column of Table 4.3 shows the item's correlation with the total score, excluding that item from the survey, also known as item-total correlation. It is a correlation between the individual item and the total score without that item. The total score is found by adding up the scores for each item. For example, for Item 1 we sum up the score for the remaining 11 items for each person. We then correlate this sum with item 1's score to get the item-total correlation. Generally, the closer the value is to 1, the better. A low item correlation indicates that the item is not measuring the same construct measured by the other items included. It is also considered that an item total correlation of less than 0.2 or 0.3 indicates that the item is not correlated very well with the scale overall and may be dropped (Everitt and Skrondal 2010). As we see from the table, all the items have item- total correlation scores above 0.3, meaning that all the items can be said to measure the same latent construct, i.e., hope. That being said, the negative items have lower values of item-total correlation, indicating that they measure hope in a slightly different way. In other words, positively worded items and the corresponding negatively worded items may not be fully equivalent.

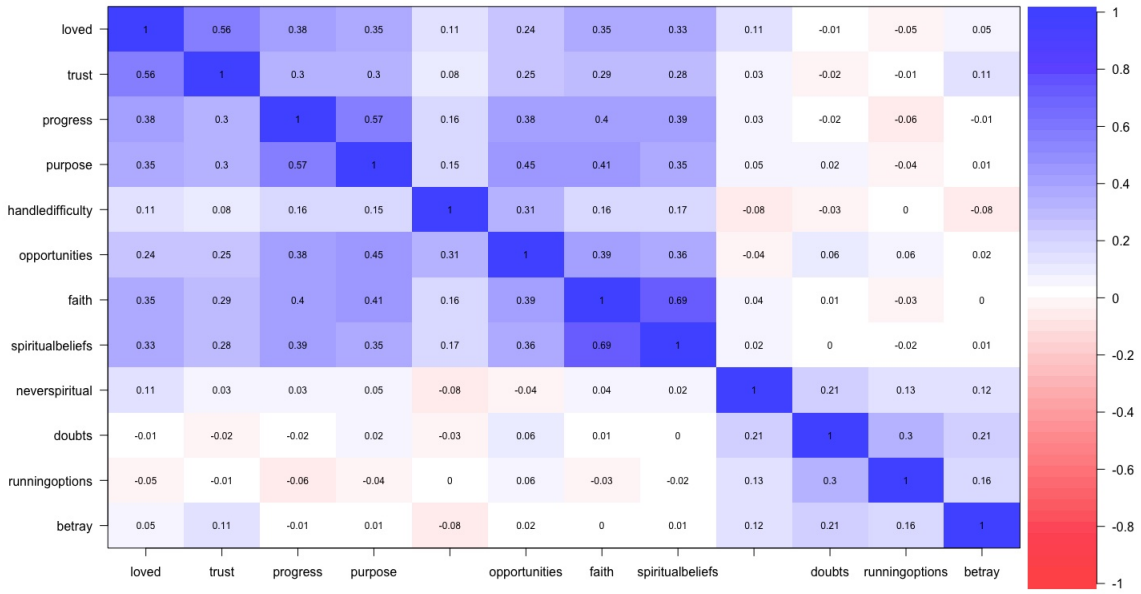


Figure 4.2: Pearson correlation for hope items

Figure 4.2 illustrates the Pearson correlation of the items graphically. Darker blue squares imply a higher correlation, whereas white and red-shaded squares indicate low or negative correlation, respectively. We see that the positive items are positively correlated with each other. This supports the idea that the positive items represent the same latent construct. The negatively worded items, namely Items 9-12, are reverse coded. Therefore, we should expect most of the squares to be blued-shaded. This would indicate that all the items represent the same construct. However, the boxes representing correlations between the positive and the negative items are mostly shaded white or red, again suggesting that the two groups are not perceived similarly by respondents. This is curious, because negative items being reverse coded, we would expect higher positive correlations between positive items and their respective negative counterparts. For example, Item 8 is a positive spirituality item and Item 9 is a negative spirituality item. If respondents strongly agreed with Item 8, we would roughly expect them to strongly disagree with Item 9 as well. With reverse coding

this would have given us a high correlation between these items. On the contrary, we see almost a zero correlation between Items 8 and 9. One possible explanation is that these items are not perceived as opposites by the respondents, and we should not treat them as positive-negative pairs.

Note also that Item 5 relating to the ability to handle difficulty is least correlated with other positive items, except with Item 6. In fact, we may notice that items relating to the same hope components are more strongly correlated with one another than with positive items from other hope components.

Chapter 5

Results

In the previous chapter, we presented summary statistics of the data from two regions and briefly talked about the hope items. In this chapter, we will present the results for both of our objectives.

5.1 Satisfying IRT Assumptions

We begin by examining the validity of the IRT assumptions. The first assumption is internal consistency which is required to make sure the hope scale contains items related only to the hope concept. Internal consistency measures whether several items that propose to measure the same construct produce similar scores. For example, if a respondent endorses a statement "I like riding bicycles" and disagrees with the statement "I hate bicycles", this indicates a good internal consistency. Psychometric test validation techniques like Cronbach's alpha are frequently used to measure internal consistency. Cronbach's alpha determines the average correlation of its items.

Our estimates for Cronbach's alpha values are presented in Table 5.1. Acceptable values of alpha are considered to range between 0.70 to 0.95 (Tavakol and Dennick 2011). From Table 5.1, our Cronbach's alpha estimates for positive items fall in the range of acceptable levels of internal consistency. However, the alpha value estimate for the 12 items taken together falls under the range of acceptable values. This might be the result of a small number of questions or a weak correlation. If a low alpha

is caused by a lack of association between items, for example, some items should be changed or deleted. The simplest way is to determine the correlation between each test item and the overall score test; those with low correlations (close to zero) should be removed to only keep the items that only measure hope (Tavakol and Dennick 2011). Therefore, based on the results of Cronbach alpha, we can only satisfy internal consistency for positive items taken separately, but not for the whole set of questions. Thus, we analyze the positive and negative items separately. There are examples in the literature of this being done, like Beaman et al. (2012) who say that aggregating aspirational questions into one index may not be legitimate and may confound the effects of reservation on the various spheres of aspiration. Thus, after validation tests, they also excluded one poorly correlated question from the aspirations index (Beaman et al. 2012).

Another assumption of IRT is item invariance. If the invariance assumption holds, items are understood the same way by all respondents. When this assumption does not hold, it might be because items were poorly worded or were interpreted differently by various samples. When an item's difficulty and discrimination, as defined in section 3.2, differ across subgroups even after controlling for the level of hope, it means that the invariance assumption does not hold and the item is said to exhibit differential item functioning (DIF). To recall, the difficulty is the minimum latent trait level required to correctly answer an item and discrimination is an item's ability to distinguish between individuals with low and high latent trait levels. DIF was discussed in the methods section in Chapter 3. We will conduct a DIF analysis of our items as the second objective of the thesis.

One way to test the assumption of local independence is to look at the discrimination parameter (Hambleton, Swaminathan, and Rogers 1991; Edelen and Reeve 2007). If items exhibit excessive covariation or dependence, they may have extremely high slopes (e.g.>4) in comparison to other items in the measure (Nguyen et al. 2014). To analyze the possible violations of the assumption of local independence,

	Cronbach's Alpha	95% CI
positive items	0.79	[0.78, 0.80]
negative items	0.48	[0.45, 0.50]
all items	0.64	[0.63, 0.66]

Table 5.1: Chronbach's Alpha

the discrimination parameter was evaluated.

5.1.1 Model choice and model fit

When designing our IRT study, we can assume that all the items are equally effective at discriminating between respondents, or we can assume that each item has a unique discrimination parameter. Thus, we are comparing a "constrained" model to an "unconstrained" model - and, as we will see, the difference between these models in terms of their fit to the data can be explicitly tested. To compare the model fit statistics of constrained and non-constrained models, we use an ANOVA (Analysis of Variance) that evaluates the difference between the two models. We did it separately for positive and negative items, in Table 5.2.

Fit of constrained vs non-constrained GRM for positive items of Hope with Likelihood ratio test						
	AIC	BIC	log.Lik	LRT	df	p.value
Constrained model	67678.55	67842.6	-33814.27			
Non-constrained model	66697.88	66907.86	-33316.94	994.67	7	<0.001
Fit of constrained vs non-constrained GRM for negative items of Hope with Likelihood ratio test						
	AIC	BIC	log.Lik	LRT	df	p.value
Constrained model	52133.07	52218.75	-26053.53			
Non-constrained model	51980.36	52085.82	-25974.18	158.7	3	<0.001

Table 5.2: Fit on Constrained vs Non-constrained

A significant p-value tells us that the non-constrained model, when we do not assume that each item has the same discrimination parameter is a better fit for

the data than the constrained model. Thus, we will use the GRM non-constrained model. Using the non-constrained model allows us to see which items discriminate respondents well by their hope levels, and which do not. The questions with a low discrimination ability, i.e. a flat slope, can then be removed as they poorly measure the hope level.

5.2 Objective 1: Item Response Theory for Hope

In the first objective, we investigate whether the survey questions measure hope and not something else. To this aim, we conduct GRM analysis and report results in Table 5.3 below. We separately consider positive and negative items and identify which questions discriminate respondents better according to their hope levels. Furthermore, having analyzed the location, or difficulty, of the item, we want to see if our items cover different parts of the hope scale, i.e., respondents with different hope levels are evenly represented. The parameters of discrimination and difficulty for each item of hope are presented in Table 5.3 below. Using the GRM to measure hope level, none of the discrimination factors exceeded 4.0. The discrimination parameters (a) of the items range from 0.66 to 2.69 and the first ($b1$), second ($b2$) and third ($b3$) difficulty parameters range from -4.38 to -1.78, -2.4 to 0.46, and 0.09 to 2.41, respectively. We use Baker and Kim's (2017) methodology for evaluating slope parameters. The results from this methodology hold under the logistic model for the item characteristic curve (Baker and Kim 2004). The discrimination parameter, a, equal to 0 means no ability for measuring latent trait, a between 0.01–0.04 means a very low ability, a between 0.35–0.64 means a poor ability, a between 0.65–1.34 means a moderate ability, a between 1.35–1.69 means a very high ability, and a greater than 1.70 signifies an extremely high level of ability (Baker and Kim 2004, p.33), where ability translates to hope in our case. All positive items except item five fall under high or very high discrimination, while three out of four negative items have moderate levels of discrimination. The lower discrimination parameters of negative items

indicate that they distinguish individuals less accurately than the positive ones.

	a	b1	b2	b3
	(SE)	(SE)	(SE)	(SE)
Positive statements				
Item 1. I feel loved by someone.	1.74	-3.40	-2.40	0.09
	(0.06)	(0.12)	(0.07)	(0.03)
Item 2. There are people in my life that I completely trust.	1.41	-3.53	-2.33	0.35
	(0.05)	(0.12)	(0.07)	(0.03)
Item 3. I'm making progress towards important goals	2.17	-2.91	-1.98	0.5
	(0.07)	(-0.09)	(0.05)	(0.02)
Item 4. I have a purpose in life.	2.05	-2.92	-1.60	0.72
	(0.07)	(-0.09)	(0.04)	(0.03)
Item 5. I can handle any current or future difficulties.	0.76	-3.90	-2.14	1.93
	(0.03)	(-0.17)	(0.10)	(0.10)
Item 6. The future will bring opportunities for a better life.	1.73	-2.87	-1.51	1.14
	(0.05)	(-0.08)	(0.04)	(0.03)
Item 7. My faith in a higher power gives me the strength to pursue my dreams.	2.69	-2.44	-1.66	0.44
	(0.09)	(-0.06)	(0.04)	(0.02)
Item 8. My spiritual beliefs have empowered me to succeed in life.	2.43	-2.53	-1.82	0.33
	(0.08)	(-0.07)	(0.04)	(0.02)
Negative Statements				
Item 9. I have never felt close to any kind of spiritual force or presence.	0.66	-4.38	-1.15	1.9
	(0.04)	(0.36)	(0.08)	(0.12)
Item 10. I have doubts about achieving those things that really matter to me.	1.80	-1.54	0.30	1.85
	(0.11)	(0.06)	(0.03)	(0.07)
Item 11. I'm running out of options for improving my life.	1.1	-1.78	0.46	2.41
	(0.06)	(0.07)	(0.04)	(0.1)
Item 12. I worry that someone may betray me.	0.78	-2.70	-0.18	2.31
	(0.04)	(0.14)	(0.04)	(0.12)

Table 5.3: IRT, Graded Response Model results for Hope

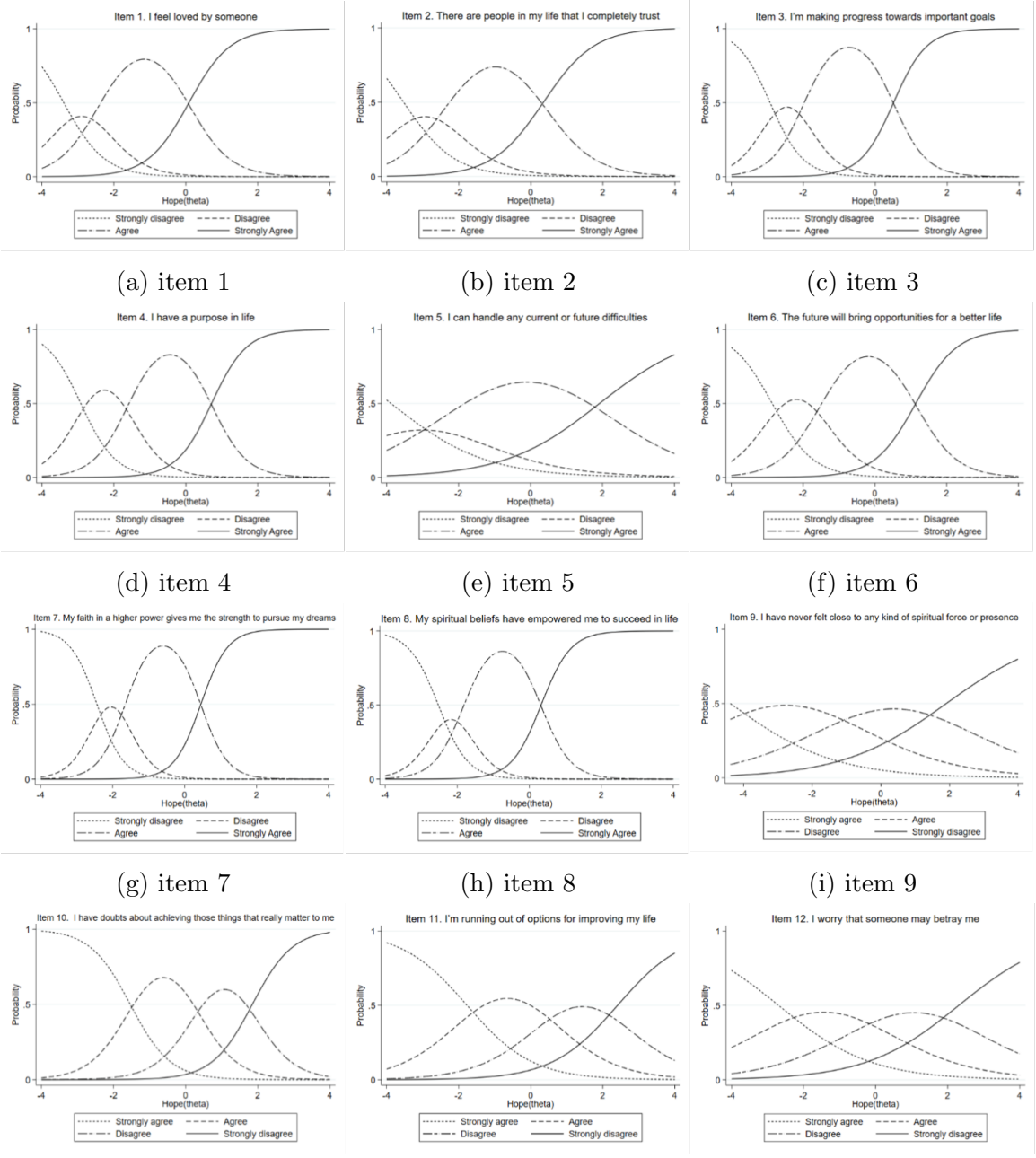
In Figure 5.1 we graph the results of the Graded Response Model for each of the 12 items that we had in Table 5.3. The vertical axis is the probability of the positive

response to the item, or the agreement with the item. The horizontal axis is the level of hope: the higher the value of the x-axis, the higher the hope level. Each graph consists of four curves, each corresponding to the level of agreement with the item, namely strongly disagree, disagree, agree, and strongly agree. Each curve shows the probability of agreeing with the item given the level of hope. For example, the solid line graph corresponds to "strongly agree". We see in 5.1(a), which is a graph for Item 1, that people with higher levels of hope were more likely to agree with Item 1 than people with low levels of hope. Ideally, we want each of the curves to form four distinct bumps. This way, each level of agreement with the item would correspond to its level of hope. We can see this in graphs for Items 1, 3, 4, 6, 7, 8, and 10. In contrast, Items 5, 9, 11, and 12 have overlapping curves, which makes it hard to map levels of agreement with the item to the hope levels. For example, a person with hope level -1 in the graph for Item 9 is equally likely to agree or disagree and strongly agree or strongly disagree with the item.

Further, we can compare positive items (Items 1-8) and negative items (Items 9-12). We can see that all positive items except Item 5 discriminate quite well. For Item 5, we see many respondents "agree" with the item regardless of their hope level. For negative items, we see that in general, the answer categories do a worse job of discriminating respondents by hope level. For example, in Items 9 and 12, given a level of hope, there are similar chances of agreeing or disagreeing with the item.

Furthermore, respondents need higher levels of hope to disagree with negative statements than to agree with positive statements. For example, Item 8 requires a low level of hope to elicit strong agreement, but Item 11 would seem to require a much higher level of hope to be endorsed. That is, even people who are only somewhat hopeful (i.e., people who have relatively low levels of hope) would likely agree with Item 8, "My spiritual beliefs have empowered them to succeed in life". In contrast, a person would probably need to be very hopeful to disagree with Item 11 "I am running out of options for improving my life". Looking at the slopes/discrimination

parameters in Table 5.3 and Figure 5.1, we can say that disagreeing with negative statements requires more effort towards increasing hope levels, compared to agreeing with positive statements. For example, on a math test, you may need less effort to increase your ability to correctly solve easy questions, while hard questions require more time and effort to reach the ability level at which you can solve them. In our case, for example, disagreeing with Item 11, "I am running out of options for improving my life", requires a substantial upgrade in hope levels compared to agreeing with Item 4 "I have a purpose in life." We also see from Figure 5.1 that in Items 5, 9, and 12 the distributions of the categories significantly overlap, meaning that individuals with very different levels of hope have a similar probability of choosing a specific category. In effect, these items do a poor job of discriminating between high and low hope individuals. Interestingly, the difficulty of negative items shown in Table 5.3 goes up to 2.41, indicating a variation among individuals with higher levels of hope, which adds substantially to positive ones. In other words, negative items were the "hardest" items for the respondents to endorse. This suggests that negative items target higher trait levels of hope than positive items.



(a) item 1 (b) item 2 (c) item 3

(d) item 4 (e) item 5 (f) item 6

(g) item 7 (h) item 8 (i) item 9

(j) item 10 (k) item 11 (l) item 12

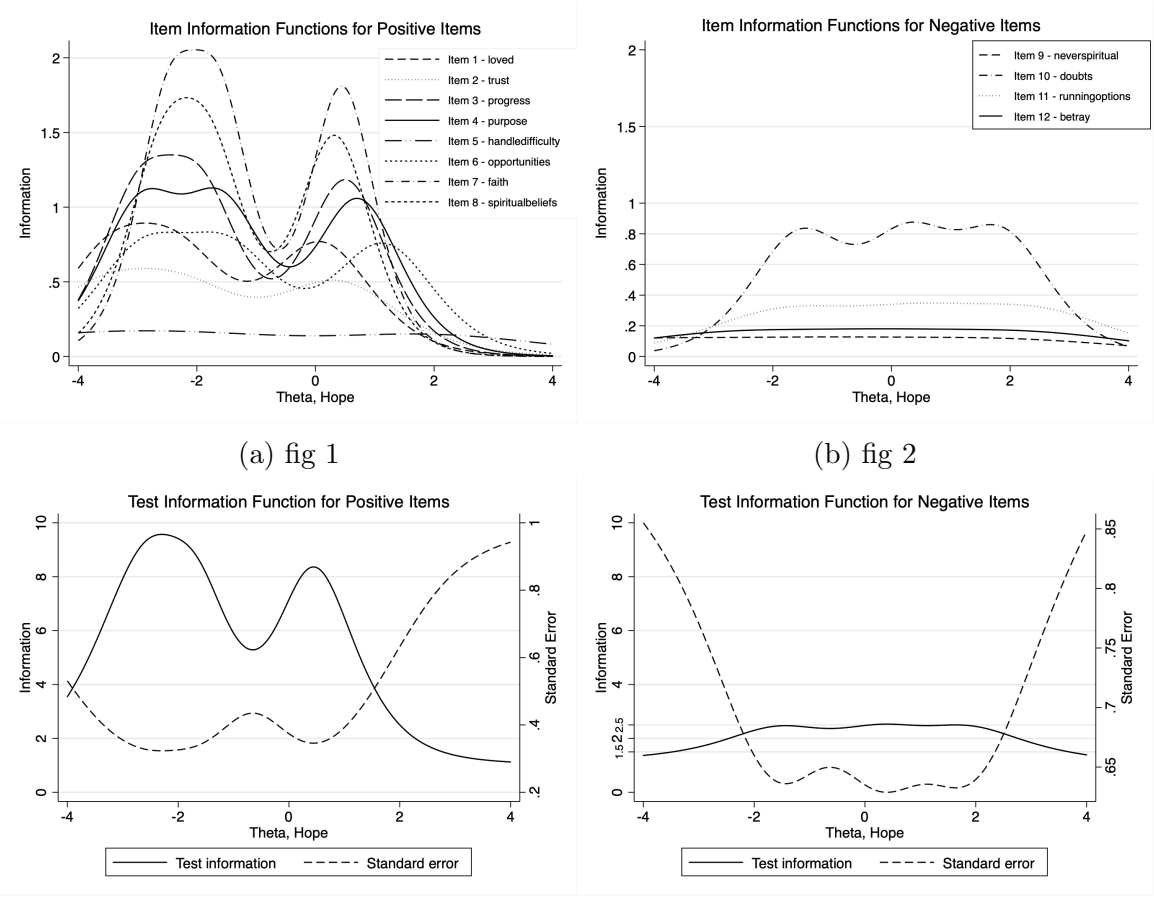
Figure 5.1: IRT for Hope

5.2.1 Item Information Functions and Test Information Functions for Hope

In this section, we consider the informativeness of the items across different levels of hope. Figures 5.2 (a) and (b) present item information curves for each positive and

negative item. The height of the curve indicates the amount of information that the item provides. The highest point on a curve represents the hope level at which the item provides the most information. In fact, an item provides the most information at a trait level that corresponds with its difficulty level, as estimated earlier. For example, Item 7, "My faith in a higher power gives me the strength to pursue my dreams", is one of the easiest items and provides the best information at a trait level of -2, which are respondents with little hope. In general, the figures show that positive items are good at measuring latent hope since they provide more information, but only at lower and average levels of hope. In other words, positive items are more capable of discriminating among individuals with lower hope levels than among people with high hope levels, presumably because most people with high hope levels agree with the statements. Figures 5.2 (c) and (d) show the graphs for the test information functions and their corresponding standard errors (SE) for the hope scale. As can be seen in this graph, the SE values were relatively low for hope values ranging from -4 to 1.5 for positive items and from -2 to 2 for negative items. Thus, positive items provide strong and precise information on hope except at high levels of hope. As for the negative items, they provide weak information which is imprecise at low and high levels of hope. Recalling our first objective which is to check whether our survey is measuring what we intend to measure, we also want to see whether it measures different levels of hope. Looking at Figures 5.2 (a) and (b), note that Items 5, 9, 11, and 12 provide very little information. It may be a good idea to replace them with items that provide more information, especially about the average and high levels of hope (θ between 1 and 3). We see that, in general, negative items are less informative, except for Item 10, which provides good information and covers people with a wider range of hope levels, especially those who have higher levels of hope. Similar findings were reported by Zanon et al. (2016), who discovered that surveys with positive and negative questions may be differently perceived. They found that in general positive statements explain latent constructs better than negative statements (Zanon et al.

2016). Positive statements are more informative at lower levels of the latent trait, while negative statements can still be useful at higher levels.



(a) fig 1

(b) fig 2

(c) fig 3

(d) fig 4

Figure 5.2: Item Information Function (IIF) and Test Information Function (TIF) of the Scioli Hope. The latent trait is shown on the horizontal axis, and the amount of information for each item for IIF and the amount of information for all items and standard error yielded by the test at any trait level are shown on the vertical axis

If we use a shortened survey with positive items, it will be better at assessing individuals with low to moderate levels of hope, but not as good in assessing individuals with higher levels of hope.

5.3 Objective 2: Differential Item Functioning Results for Hope

To examine whether items behaved equivalently across a range of different subgroups of people, we conduct differential item functioning (DIF) analyses, which also comprise a test of the invariance assumption of the IRT. In DIF analyses, six grouping criteria were examined: gender, region, age, whether respondents had access to food and water, and treatment. We divided them into two groups, focal and reference, as explained in Section 3.4: the "focal" group is the one suspected to be at a disadvantage to the "reference" group. We divide the sample into reference and focal groups, respectively, like so: males vs females, adults vs adolescents (10-19 years old), Kigoma vs Iringa; those who for the past one year met their daily food needs, and those who did not; those who have access to protected water or have tap water at home versus those who use unprotected sources of water. We also consider the DIF for the control and intervention groups, as defined in Section 3.4. DIF analyses compare item endorsement rates in the focal group compared to the reference group, conditional on the latent trait. For example, we would compare respondents who have no access to protected water (focal group) compared to the rest who have access to any type of protected water (reference group) given the same level of the trait. When respondents with the same overall hope score have differing probabilities of endorsing an individual item, the item is said to demonstrate DIF; in other words, the item acts differentially between the two groups.

	Gender (SE) (1-male; 2-female)	Region (SE) (0-Kigoma; 1-Iringa)	Age (SE) (0-adolecents (10-19 years); 1-adult;)	Food Security (SE) (0-secure, 1-no)	Access to protected water (SE) (0-yes, 1-no)	Treatment (SE) (0-control; 1-intervention
Item 1. I feel loved by someone.	-0.04 (0.07)	-0.62 (0.08)***	-0.24 (0.07)**	0.02 (0.09)	0.12 (0.08)	0.05 (0.03)
Item 2. There are people in my life that I completely trust.	-0.1 (0.06)	-0.64 (0.07)***	-0.51 (0.07)***	-0.14 (0.08)*	0.06 (0.07)	0.07 (0.03)**
Item 3. I'm making progress towards important goals	-0.18 (0.08)**	0.09 (0.08)	-0.04 (0.08)	0.08 (0.1)	-0.20 (0.08)**	-0.04 (0.04)
Item 4. I have a purpose in life.	-0.09 (0.07)	0.58 (0.08)***	0.19 (0.07)**	0.22 (0.09)**	0.15 (0.08)**	-0.04 (0.25)
Item 5. I can handle any current or future difficulties.	-0.25 (0.06)***	-0.15 (0.06)**	1.14 (0.06)***	-0.26 (0.07)***	-0.03 (0.06)	-0.03 (0.03)
Item 6. The future will bring opportunities for a better life.	-0.17 (0.07)**	0.8 (0.07)***	0.08 (0.07)	0.13 (0.09)	0.01 (0.07)	-0.01 (0.03)
Item 7. My faith in a higher power gives me the strength to pursue my dreams.	0.15 (0.08)*	0.13 (0.12)	1.14 (0.08)*	-0.3 (0.1)	0.18 (0.09)**	0.02 (0.04)
Item 8. My spiritual beliefs have empowered me to succeed in life.	0.25 (0.08)**	-0.11 (0.08)	-0.15 (0.08)	-0.32 (0.1)**	-0.01 (0.08)	-0.03 (0.04)
Item 9. I have never felt close to any kind of spiritual force or presence.	0.07 (0.05)	0.09 (0.06)*	0.09 (0.05)*	0.09 (0.07)	-0.04 (0.06)	-0.04 (0.03)
Item 10. I have doubts about achieving those things that really matter to me.	0.01 (0.05)	0.27 (0.06)***	0.16 (0.05)**	-0.45 (0.07)***	0.04 (0.06)	-0.00 (0.03)
Item 11. I'm running out of options for improving my life.	0.03 (0.05)	0.42 (0.06)***	-0.37 (0.05)***	-0.38 (0.07)***	-0.14 (0.06)**	0.04 (0.03)*
Item 12. I worry that someone may betray me.	0.04 (0.05)	-0.15 (0.06)**	-0.27 (0.05)***	-0.17 (0.65)**	0.02 (0.06)	-0.02 (0.03)

*Note: *=p<0.1; ** p<0.05;***= p<0.01*
Food security: In the past 12 months how did your household meet daily food needs?
Water access: What is the main water source for your household?

Table 5.4: DIF for gender, age, region, food security, access to water and treatment for Hope

The absence of DIF indicates that the probability of a person's response to these items depends only on her or his hope level and not on other factors. In the context of the measurement of hope for females and males, a person's response to an item will depend only on the level of hope she or he has and not on whether the person is female or male.

Table 5.4 presents DIF analyses of the survey items conditional on hope. From the table, we see that there is no DIF for Items 3, 4, 6, and 7 among adults and adolescents, for Items 3, 7, 8, 9, and 12 among people living in Kigoma and Iringa, for Items 3, 4, 7, and 9 among people who could meet their daily food needs and those who could not, for Items 4, 5, 7, 8, 9, 10, and 12 among those who have access to protected water or have tap water at home and those who access unprotected sources of water, and for 10 items for the treatment group.

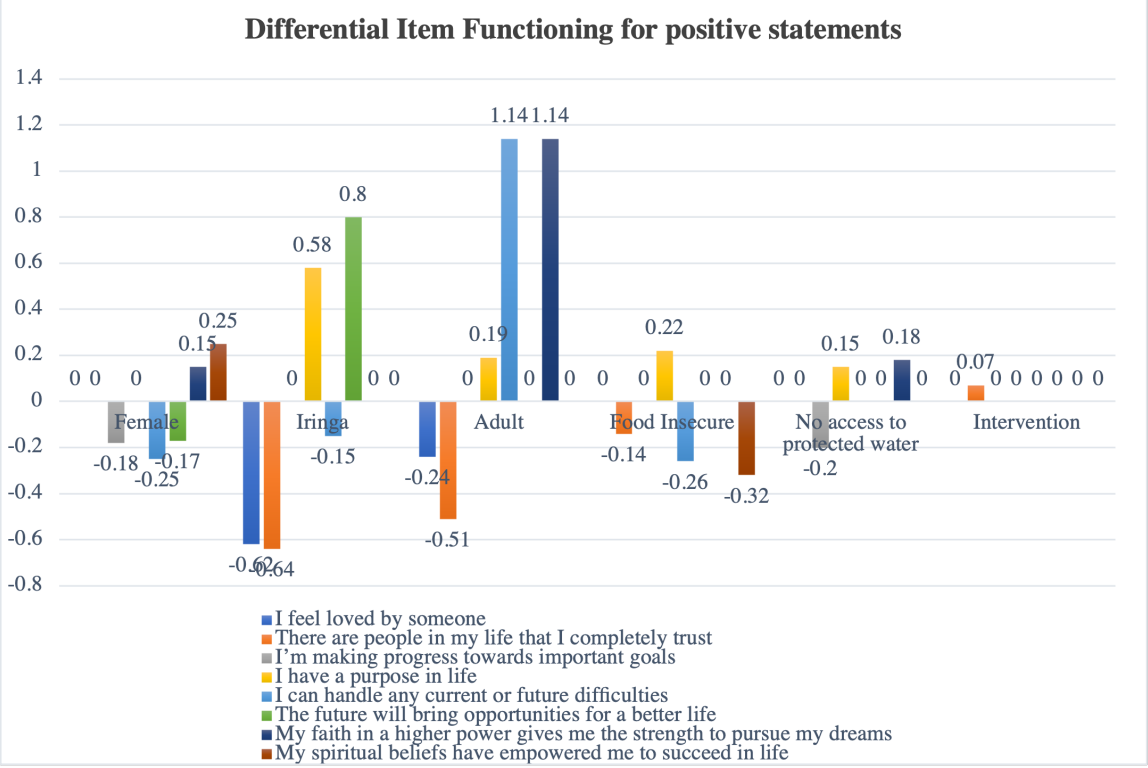


Figure 5.3: DIF for positive items with significant DIF

At the same time, the results demonstrate some DIF for items by groups. To make the results visually clearer we present Figure 5.3 which includes only positive items with statistically significant DIF. The same is done for negative items in Figure 5.4. From the figure we see that females, given the same hope level as males, are less likely to agree that they are making progress, that they can handle current and future difficulties, and that the future will bring opportunities for a better life. However,

they have a stronger belief in faith and spirituality that will bring them closer to their dreams and success compared to males with the same hope levels. Adults are more likely to agree that they can handle difficulties and that their faith in a higher power gives them the strength to pursue their dreams. In other words, adults have 141% probability of agreeing or strongly agreeing with these items that adolescents. However, adults with the same level of hope as adolescents are 24% and 51% less likely to agree that they feel loved and have people whom they can trust. Moreover, DIF analyses suggest that people who do not have access to protected water are less likely to agree with the statement "I'm making progress towards important goals."However, those who did not manage to meet their daily food needs for the last year and have no access to water are more likely to say that they have a purpose in life, even though they have equal hope levels with those who are food secure and have access to protected water. There is also DIF by region, as seen in Figure 5.3. People living in Iringa, given the same level of hope as people from Kigoma, are 62% and 64% less likely to agree that they feel loved by someone and have people whom they can completely trust. However, in Iringa, people are 58% more likely to have a purpose in life and 80% are more likely to believe that the future will bring opportunities for a better life.

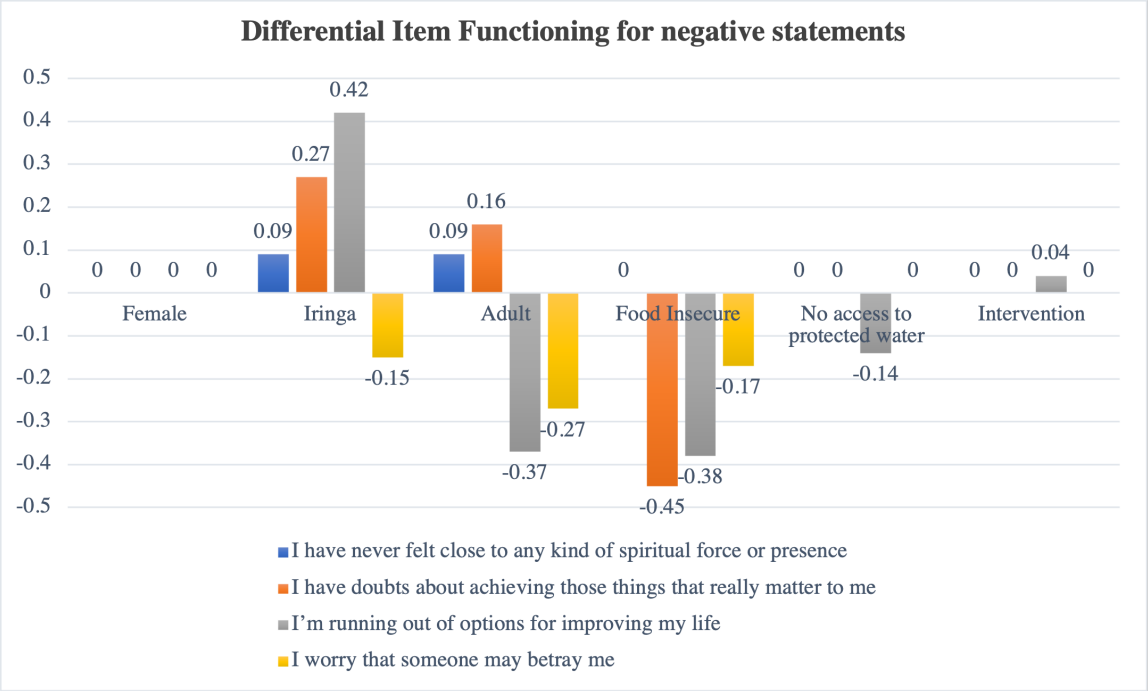


Figure 5.4: DIF for negative items with significant DIF

Figure 5.4 visually presents negatively stated items with significant DIF. Results suggest that people living in Iringa are more likely to disagree that they have never felt close to any kind of spiritual force or have doubts about achieving goals or running out of options for improving life, while they worry more that someone may betray them. Also, adults are less likely to disagree that they are running out of options and that someone may betray them. Those who did not manage to meet their daily food needs for the last year were also less likely to disagree with three negative statements.

Chapter 6

Conclusions

Empirical economists emphasize the importance of comparing apples to apples, i.e., holding all else constant as the effect of an intervention is assessed. The increasing use of randomized experiments ushered in by the credibility revolution in economics brought us closer to making such adequate comparisons. However, as economists venture into studying the effects of latent constructs like hope, which were traditionally the domain of other social sciences, they inadvertently make apples-to-oranges comparisons by not using proper measurement tools. On the surface, the word "hope" has the same meaning to different people, but in reality, its understanding varies with culture, geography, upbringing, and many other factors. While psychologists and sociologists are wary of these discrepancies and use statistical techniques to keep their measurement tools in check, economists are just beginning to appreciate the complexity of measuring latent constructs.

This thesis analyzed the baseline survey data collected in Tanzania to appraise the robustness of our survey items to measurement errors to make necessary corrections for the further stages of the study involving randomized experiments. The hope survey was adapted from an existing hope scale typically used in developed countries, and thus its ability to measure hope needed to be tested in the context of Tanzania first before proceeding with the later stages of the study. The first objective of the thesis was to identify if survey items can successfully distinguish respondents by hope

levels. The second objective was to see whether survey items were interpreted the same way or differently across different population categories.

The first objective was addressed in Section 5.2. We find that our survey items indeed can successfully discriminate among survey participants by their levels of hope. Using the methodology outlined in Baker and Kim (2004) we find that all survey items have at least a moderate ability to discriminate respondents by hope level. At the same time, positive items discriminate better than negative items. In fact, all positive items except Item 5 demonstrated a high or very high discrimination ability, while three out of four negative items had moderate levels of discrimination, and the fourth had a very high discrimination ability. Our findings show that people need higher levels of hope to disagree with them rather than agreeing with positive items. It means that negative items were the hardest items for our respondents. They provide less information than positive items but cover people with a wider range of hope levels, especially those who have higher levels of hope. It means that we can differentiate people at low and average levels of hope but not good enough at higher hope levels. Thus, we suggest that future research include more items that are informative at higher hope levels.

The second objective was addressed in Chapter 5.3. Overall, we find that there are differences in how groups of respondents answered the survey. First of all, we see almost no DIF for the treatment and control groups in the measurement of hope, which means that the groups are comparable enough for conducting RCTs in the future phases of the study. However, we find high DIF in the responses of adults and adolescents. If this DIF is ignored in the future stages of the study, adolescents will be erroneously scored as less hopeful than adults. This is because of two items in our study with which adults would likely agree with if they were hopeful, but agreeing with these items would be indicative of a high hope level for adolescents. One way to address this problem is to remove the items contributing to the high DIF. This is not necessarily the best procedure because item removal can change the meaning of the

construct and lowered reliability (Teresi 2006; Hambleton 2006), i.e. the remaining items may not be enough to encapsulate hope. Moreover, some of our items show DIF cancellation, a situation when DIFs cancel each other out, and their removal could bias the results. This happens when we compare Iringa and Kigoma for most of the items. An alternative solution is to compare results within groups and not across groups.

Bibliography

Adams V.H., Snyder C.R. Rand K.L. King E.A. Sigman D.R. and K.M. Pulvers (2002). *Hope in the Workplace. In Giacolone, R. and Jurkiewicz, C., Eds., Handbook of Workplace Spirituality and Organization Performance, Sharpe, New York.*

Anderson, Craig L. and David B. Feldman (2020). "Hope and Physical Exercise: The Contributions of Hope, Self-Efficacy, and Optimism in Accounting for Variance in Exercise Frequency". In: *Psychological reports* 123.4, pp. 1145–1159. ISSN: 1558-691X. DOI: 10.1177/0033294119851798. URL: https://pubmed.ncbi.nlm.nih.gov/31142190/.

Appadurai, Arjun (2004). "Culture and Public Action. In Rao, Vijayendra and Michael Walton (ed). The Capacity to Aspire: Culture and the Terms of Recognition". In: *Culture*, pp. 59–84.

Arampatzi, Efstratia et al. (2018). "Unhappy development: Dissatisfaction with Life on the Eve of the Arab Spring". In: *Review of Income and Wealth* 64, S80–S113.

Aspinwall, Lisa G. and Samantha L. Leaf (2002). "In search of the unique aspects of hope: Pinning our hopes on positive emotions, future-oriented thinking, hard times, and other people". In: *Psychological Inquiry* 13.4, pp. 276–288. ISSN: 1047840X. DOI: 10.1207/S15327965PLI1304.

Baker, Frank B and Seock-Ho Kim (2004). *Item response theory: Parameter estimation techniques*. CRC press.

Bandura, Albert (1982). "Self-efficacy mechanism in human agency". In: *American Psychologist* 37.2, pp. 122–147. ISSN: 0003066X. DOI: 10.1037/0003-066X.37.2.122.

Bandura, Albert et al. (2019). "Leading an African Renaissance". In: *Leading an African Renaissance* 84.2, pp. 133–154. ISSN: 00219010. DOI: 10.1037/1099-9809.13.2.94. URL: https://doi.org/10.1080/13674676.2019.1594740https://via.library.depaul.edu/csh_etd/296http://dx.doi.org/10.1016/j.clinthera.2014.04.006.

Banerjee, Abhijit et al. (2015). "A multifaceted program causes lasting progress for the very poor: Evidence from six countries". In: *Science* 348.6236. ISSN: 10959203. DOI: 10.1126/science.1260799.

Bar-Tal, Daniel (2001). "Why does fear override hope in societies engulfed by intractable conflict, as it does in the Israeli society?" In: *Political Psychology* 22.3, pp. 601–627.

Beaman, Lori et al. (2012). "Female leadership raises aspirations and educational attainment for girls: A policy experiment in India". In: *Science* 335.6068, pp. 582–586. ISSN: 10959203. DOI: 10.1126/science.1212382.

Bénabou, Roland et al. (n.d.). "Self-knowledge and self-regulation: An economic approach". In: *Citeseer* (). URL: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.448.5005&rep=rep1&type=pdf.

Benzein, E. and B. i. Saveman (1998). "One step towards the understanding of hope: a concept analysis". In: *International Journal of Nursing Studies* 35.6, pp. 322–329. ISSN: 0020-7489. DOI: https://doi.org/10.1016/S0020-7489(98)00045-5. URL: https://www.sciencedirect.com/science/article/pii/S0020748998000455.

Berg, Carla J. et al. (2011). "The role of hope in engaging in healthy behaviors among college students". In: *American Journal of Health Behavior* 35.4, pp. 402–415. ISSN: 10873244. DOI: 10.5993/AJHB.35.4.3.

Besley, Timothy (2017). "Aspirations and the political economy of inequality". In: *Oxford Economic Papers* 69.1, pp. 1–35. ISSN: 14643812. DOI: 10.1093/oep/gpw055.

Birnbaum, A. (1968). "Some Latent Trait Models and Their Use in Inferring an Examinee's Ability. Statistical Theories of Mental Test Scores". In.

Bloem, Jeffrey and Bruce Wydick (2021). "All I Really Need to Know I Learned In Kindergarten? Evidence from the Philippines". In: *Economic Development and Cultural Change*. ISSN: 0013-0079. DOI: 10.1086/715502.

Bloem, Jeffrey R. et al. (2018). "Measuring Hope: A Quantitative Approach with Validation in Rural Myanmar". In: *Journal of Development Studies* 54.11, pp. 2078–2094. ISSN: 17439140. DOI: 10.1080/00220388.2017.1385764. URL: https://doi.org/10.1080/00220388.2017.1385764.

Braaten, Leif J. (1991). "Group cohesion: A new multidimensional model". In: *Group* 15.1, pp. 39–55. ISSN: 03624021. DOI: 10.1007/BF01419845.

Briggs, Charles L (1986). *Learning how to ask: A sociolinguistic appraisal of the role of the interview in social science research*. 1. Cambridge university press.

Brislin, Richard W. (1970). "Back-Translation for Cross-Cultural Research:" in: *http://dx.doi.org/10.1177/135910457000100301* 1.3, pp. 185–216. ISSN: 15525422. DOI: 10.1177/135910457000100301. URL: https://journals.sagepub.com/doi/10.1177/135910457000100301.

Bryan, Gharad, James J Choi, and Dean Karlan (June 2020). "Randomizing Religion: the Impact of Protestant Evangelism on Economic Outcomes*". In: *The Quarterly Journal of Economics* 136.1, pp. 293–380. ISSN: 0033-5533. DOI: 10.1093/qje/qjaa023. eprint: https://academic.oup.com/qje/article-pdf/136/1/293/35073187/qjaa023.pdf. URL: https://doi.org/10.1093/qje/qjaa023.

Camilli, Gregory (2006). "Test fairness". In: *Educational measurement* 4, pp. 221–256.

CCNMTL (2003). *Quantitative Method in the Social Sciences e-lessons*. Available at https://ccnmtl.columbia.edu/projects/qmss/images/target.gif (2022/03/07).

Chan, Joseph, Ho Pong To, and Elaine Chan (2006). *Reconsidering social cohesion: Developing a definition and analytical framework for empirical research*. DOI: 10.1007/s11205-005-2118-1.

Chen, Fang Fang et al. (2012). "Modeling general and specific variance in multifaceted constructs: A comparison of the bifactor model to other approaches". In: *Journal*

of Personality 80.1, pp. 219–251. ISSN: 00223506. DOI: 10.1111/j.1467-6494.2011.00739.x.

Chen, Gilad, Stanley M. Gully, and Dov Eden (2001). "Validation of a New General Self-Efficacy Scale". In: *https://doi.org/10.1177/109442810141004* 1, pp. 62–83. ISSN: 10944281. DOI: 10.1177/109442810141004.

Core Guide: Measure Development, (2020). Tech. rep. URL: https://sites.globalhealth.duke.edu/rdac/wp-content/uploads/sites/27/2020/08/Core-Guide_Measure-Design-05-04-20.pdf.

Darolia, Rajeev and Bruce Wydick (2011). "The Economics of Parenting, Self-esteem and Academic PerformanceTheory and a Test". In: *Economica* 78.310, pp. 215–239. ISSN: 00130427. DOI: 10.1111/j.1468-0335.2009.00802.x.

De Sá, Antonio Reis et al. (2019). "Can gender and age impact on response pattern of depressive symptoms among college students? A differential item functioning analysis". In: *Frontiers in Psychiatry* 10.FEB. ISSN: 16640640. DOI: 10.3389/fpsyt.2019.00050.

Dorans, Neil J. and Neal M. Kingston (1985). "the Effects of Violations of Unidimensionality on the Estimation of Item and Ability Parameters and on Item Response Theory Equating of the Gre Verbal Scale". In: *Journal of Educational Measurement* 22.4, pp. 249–262. ISSN: 17453984. DOI: 10.1111/j.1745-3984.1985.tb01062.x.

Drahos, Peter (2004). "Trading in public hope". In: *The Annals of the American Academy of Political and Social Science* 592.1, pp. 18–38.

Drasgow, Fritz and Charles K. Parsons (1983). "Application of Unidimensional Item Response Theory Models to Multidimensional Data". In: *Applied Psychological Measurement* 7.2, pp. 189–199. ISSN: 15523497. DOI: 10.1177/014662168300700207.

Du, Hongfei and Ronnel B. King (2013). "Placing hope in self and others: Exploring the relationships among self-construals, locus of hope, and adjustment". In: *Personality and Individual Differences* 54.3, pp. 332–337. ISSN: 01918869. DOI: 10.1016/J.PAID.2012.09.015.

Duflo, Esther (2012). *Lack of Hope and Persistence of Poverty.* URL: https://www.econ.cam.ac.uk/events-files/seminars/Marshall_Lecture/2012-13/hope_cambridge.pdf.

Duggal, Devika, Amanda Sacks-Zimmerman, and Taylor Liberta (2016). "The Impact of Hope and Resilience on Multiple Factors in Neurosurgical Patients". In: *Cureus* 8.10. ISSN: 2168-8184. DOI: 10.7759/cureus.849.

Dunn, Thomas J., Thom Baguley, and Vivienne Brunsden (2014). "From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation". In: *British Journal of Psychology* 105.3, pp. 399–412. ISSN: 20448295. DOI: 10.1111/BJOP.12046.

Easterlin, Richard A (2001). "Income and happiness: Towards a unified theory". In: *The economic journal* 111.473, pp. 465–484.

Edelen, Maria Orlando and Bryce B. Reeve (2007). "Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement". In: *Quality of Life Research* 16.SUPPL. 1, pp. 5–18. ISSN: 09629343. DOI: 10.1007/S11136-007-9198-0.

Ekici, Tufan and Selda Koydemir (2016). "Income expectations and happiness: evidence from British panel data". In: *Applied Research in Quality of Life* 11.2, pp. 539–552.

Embretson S. E., Reise S. P. (2000). "Item response theory for psychologists". In: URL: https://psycnet.apa.org/record/2000-03918-000.

Everitt, Brian S and Anders Skrondal (2010). "The Cambridge dictionary of statistics". In.

Franko, Debra L et al. (2008). "Motivation, self-efficacy, physical activity and nutrition in college students: Randomized controlled trial of an internet-based education program". In: *Preventive medicine* 47.4, pp. 369–377.

Furr, R. Michael (2014). "Scale Construction and Psychometrics for Social and Personality Psychology". In: *Scale Construction and Psychometrics for Social and Personality Psychology*. DOI: 10.4135/9781446287866.

Gallagher, Matthew W. and Shane J. Lopez (2009). "Positive expectancies and mental health: Identifying the unique contributions of hope and optimism". In: *Journal of Positive Psychology* 4.6, pp. 548–556. ISSN: 17439760. DOI: 10.1080/17439760903157166.

Genç, Emel and Gökmen Arslan (2021). "Optimism and dispositional hope to promote college students' subjective well-being in the context of the COVID-19 pandemic". In: *Journal of Positive School Psychology* 5.2, pp. 87–96. ISSN: 27177564. DOI: 10.47602/JPSP.V5I2.255.

Genicot, Garance and Debraj Ray (2017). "Aspirations and Inequality". In: *Econometrica* 85.2, pp. 489–519. ISSN: 0012-9682. DOI: 10.3982/ecta13865.

Glewwe, Paul, Phillip H Ross, and Bruce Wydick (2014). "Developing Aspirations: The Impact of Child Sponsorship on Self-Esteem and Life Expectations". In: *University of San Francisco*, pp. 1–34.

Gorsuch, Richard L (1983). *Factor analysis . Hillsdale, New Jersey: L.*

Hambleton, RK, H Swaminathan, and HJ Rogers (1991). *Fundamentals of item response theory*.

Hambleton, Ronald K (2006). "Good practices for identifying differential item functioning". In: *Medical care* 44.11, S182–S188.

Herth, Kaye (1992). "Abbreviated instrument to measure hope: development and psychometric evaluation". In: *Journal of Advanced Nursing* 17.10, pp. 1251–1259. ISSN: 13652648. DOI: 10.1111/j.1365-2648.1992.tb01843.x.

Hutz, Claudio S et al. (2014). "The Relationship of Hope, Optimism, Personality in Brazilians and Americans". In: *Psychology* 5.April, pp. 514–522.

IRIS, (2017). Tech. rep. URL: https://www.iris-france.org/wp-content/uploads/2017/03/Tanzania-Final-report.pdf.

Janzen, Sarah A. et al. (2017). "Aspirations failure and formation in rural Nepal". In: *Journal of Economic Behavior Organization* 139, pp. 1–25. ISSN: 0167-2681. DOI: 10.1016/J.JEBO.2017.04.003.

Jeannotte, M. Sharon (2003). "Singing alone? The contribution of cultural capital to social cohesion and sustainable communities". In: *International Journal of Cultural Policy* 9.1, pp. 35–49. ISSN: 14772833. DOI: 10.1080/1028663032000089507.

Jensen, Robert and Emily Oster (2009). "The power of TV: Cable television and women's status in India". In: *Quarterly Journal of Economics* 124.3, pp. 1057–1094. ISSN: 00335533. DOI: 10.1162/qjec.2009.124.3.1057.

Jerusalem M., Schwarzer R. (1992). *Self-efficacy as a resource factor in stress appraisal processes*. URL: https://psycnet.apa.org/record/1992-97719-009.

Klein, Josh R. (2014). "Durkheim, Emile". In: *The Encyclopedia of Theoretical Criminology*. John Wiley Sons, Ltd, pp. 1–5. DOI: 10.1002/9781118517390.wbetc086. URL: https://onlinelibrary.wiley.com/doi/10.1002/9781118517390.wbetc086.

Kleist, Nauja and Stef Jansen (2016a). "Introduction: Hope over Time—Crisis, Immobility and Future-Making". In: *History and Anthropology* 27.4, pp. 373–392. ISSN: 02757206. DOI: 10.1080/02757206.2016.1207636.

Kleist, Nauja and Stef Jansen (2016b). "Introduction: Hope over Time—Crisis, Immobility and Future-Making". In: *History and Anthropology* 27.4, pp. 373–392. ISSN: 02757206. DOI: 10.1080/02757206.2016.1207636.

Langer, Arnim et al. (2017). "Conceptualising and Measuring Social Cohesion in Africa: Towards a Perceptions-Based Index". In: *Social Indicators Research* 131.1, pp. 321–343. ISSN: 15730921. DOI: 10.1007/s11205-016-1250-4. URL: https://lirias.kuleuven.be/2790378?limo=0.

Luszczynska, Aleksandra, Urte Scholz, and Ralf Schwarzer (2005). "The general self-efficacy scale: Multicultural validation studies". In: *Journal of Psychology: Interdisciplinary and Applied* 139.5, pp. 439–457. ISSN: 19401019. DOI: 10.3200/JRLP.139.5.439-457.

Luthans, Fred and Susan M. Jensen (2002). "Hope: A New Positive Strength for Human Resource Development". In: *Human Resource Development Review* 1.3, pp. 304–322. ISSN: 15344843. DOI: 10.1177/1534484302013003.

Lybbert, Travis J. and Bruce Wydick (2018). "Poverty, aspirations, and the economics of hope". In: *Economic Development and Cultural Change* 66.4, pp. 709–753. ISSN: 00130079. DOI: 10.1086/696968.

Lybbert, Travis J et al. (2016). "Poverty, Aspirations, and the Economics of Hope: A Framework for Study with Preliminary Results from the Oaxaca Hope Project". In.

Marlon, Jennifer R. et al. (2019). "How Hope and Doubt Affect Climate Change Mobilization". In: *Frontiers in Communication* 4.May, pp. 1–15. ISSN: 2297900X. DOI: 10.3389/fcomm.2019.00020.

Nathan, Dev (2005). "Capabilities and aspirations". In: *Economic and political weekly*, pp. 36–40.

National Bureau of Statistics (2014). Tech. rep. URL: https://www.tanzania.go.tz/egov_uploads/documents/TANZANIA_MAINLAND_SOCIO_ECONOMIC_PROFILE_sw.pdf.

Neumark-Sztainer, Dianne et al. (2003). "Correlates of fruit and vegetable intake among adolescents: Findings from Project EAT". In: *Preventive medicine* 37.3, pp. 198–208.

Nguyen, Tam H. et al. (2014). *An introduction to item response theory for patient-reported outcome measurement*. DOI: 10.1007/s40271-013-0041-0. URL: /pmc/

articles/PMC4520411//pmc/articles/PMC4520411/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC4520411/.

Nothwehr, F., D. O. Clark, and A. Perkins (2013). "Hope and the Use of Behavioral Strategies Related to Diet and Physical Activity". In: *Journal of human nutrition and dietetics : the official journal of the British Dietetic Association* 26.0 1, p. 159. ISSN: 09523871. DOI: 10.1111/JHN.12057.

Packer, C. et al. (2020). "Hope, Self-Efficacy, and Crushed Dreams: Exploring How Adolescent Girls' Future Aspirations Relate to Marriage and Childbearing in Rural Mozambique". In: *Journal of Adolescent Research* 35.5, pp. 579–604. DOI: 10.1177/0743558419897385.

Ray, Debraj (2006). "Aspirations, poverty, and economic change". In: *Understanding Poverty* March. DOI: 10.1093/0195305191.003.0028.

Redlich-Amirav, Dorit et al. (2018). "Psychometric properties of Hope Scales: A systematic review". In: *International journal of clinical practice* 72.7. ISSN: 1742-1241. DOI: 10.1111/IJCP.13213. URL: https://pubmed.ncbi.nlm.nih.gov/29920883/.

Reese, L. M. (1999). "A Classical Test Theory Perspective on LSAT Local Item Dependence. LSAC Research Report Series. Statistical Report". In.

Reise, Steven P. et al. (2013). "Multidimensionality and Structural Coefficient Bias in Structural Equation Modeling: A Bifactor Perspective". In: *Educational and Psychological Measurement* 73.1, pp. 5–26. ISSN: 15523888. DOI: 10.1177/0013164412449831.

Rustøen, Tone, Bruce A. Cooper, and Christine Miaskowski (2011). "A longitudinal study of the effects of a hope intervention on levels of hope and psychological distress in a community-based sample of oncology patients". In: *European Journal of Oncology Nursing* 15.4, pp. 351–357. ISSN: 14623889. DOI: 10.1016/j.ejon.2010.09.001. URL: http://dx.doi.org/10.1016/j.ejon.2010.09.001.

Samejima, Fumiko (1969). "Estimation of latent ability using a response pattern of graded scores". In: *Psychometrika* 34.S1, pp. 1–97. ISSN: 0033-3123. DOI: 10.1007/BF03372160.

Schrank, Beate et al. (2011). "Development and validation of an integrative scale to assess hope". In: *Health expectations : an international journal of public participation in health care and health policy* 14.4, pp. 417–428. ISSN: 1369-7625. DOI: 10.1111/J.1369-7625.2010.00645.X. URL: https://pubmed.ncbi.nlm.nih.gov/21122040/.

Schweizer, Karl (2011). "On the changing role of Cronbach's α in the evaluation of the quality of a measure". In: *European Journal of Psychological Assessment* 27.3, pp. 143–144. ISSN: 10155759. DOI: 10.1027/1015-5759/A000069.

Scioli, A and H Biller (2009). *Hope in the age of anxiety*.

Scioli, A et al. (2011). "Hope: Its nature and measurement". In: *Psychology of Religion and Spirituality* 3.2, pp. 78–97. ISSN: 19411022. DOI: 10.1037/a0020903.

Sen, Amartya (2014). "Development as freedom (1999)". In: *The globalization and development reader: Perspectives on development and global change* 525.

Sherer, Mark et al. (1982). "The Self-Efficacy Scale: Construction and Validation:" in: *http://dx.doi.org/10.2466/pr0.1982.51.2.663* 51.2, pp. 663–671. ISSN: 0033-2941.

DOI: 10.2466/PR0.1982.51.2.663. URL: https://journals.sagepub.com/doi/10.2466/pr0.1982.51.2.663.

Sibbald, B and M Roland (1994). "Understanding controlled trials. Why are randomised controlled trials important?" In: *ncbi.nlm.nih.gov* 165, pp. 137–178. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2665449/.

Snyder, C. R. (2002). "Hope theory: Rainbows in the mind". In: *Psychological Inquiry* 13.4, pp. 249–275. ISSN: 1047840X. DOI: 10.1207/S15327965PLI1304_01.

Snyder, C. R. et al. (1991). "The will and the ways: Development and validation of an individual-differences measure of hope." In: *Journal of Personality and Social Psychology* 60.4, pp. 570–585. ISSN: 0022-3514. DOI: 10.1037//0022-3514.60.4.570.

Snyder, CR et al. (2000). "Hope theory: Updating a common process for psychological change." In: URL: https://psycnet.apa.org/record/2001-00353-007.

Steiger, James H. (1990). "Structural Model Evaluation and Modification: An Interval Estimation Approach". In: *Multivariate Behavioral Research* 25.2, pp. 173–180. ISSN: 15327906. DOI: 10.1207/s15327906mbr2502_4.

Stevenson, KT and N Peterson (2016). *Motivating action through fostering climate change hope and concern and avoiding despair among adolescents. Sustainability (Switzerland) 8: 1–10.*

Talcott, Parsons (2013). *The social system.* Routledge.

Tanguy, Bernard et al. (2014). "The Future in Mind: Aspirations and Forward-Looking Behaviour in Rural Ethiopia". In: *SSRN Electronic Journal* September. ISSN: 1556-5068. DOI: 10.2139/ssrn.2514590.

Tanzania, Girls Not Brides, (2021). Tech. rep. URL: https://atlas.girlsnotbrides.org/map/tanzania/.

Tanzania HIV Impact Survey 2016-2017: Final Report, (2018). Tech. rep. Dar es Salaam. URL: https://phia.icap.columbia.edu/wp-content/uploads/2019/06/FINAL_THIS-2016-2017_Final-Report__06.21.19_for-web_TS.pdf.

Tanzania National Nutrition Survey, (2018). Tech. rep. URL: https://www.unicef.org/tanzania/reports/tanzania-national-nutrition-survey-2018.

Tavakol, Mohsen and Reg Dennick (2011). "Making sense of Cronbach's alpha". In: *International journal of medical education* 2, pp. 53–55. ISSN: 20426372. DOI: 10.5116/ijme.4dfb.8dfd.

Tennen Howard, Affleck Glenn and Ruth Tennen (2002). "Clipped feathers: The theory and measurement of hope". In: *Psychological Inquiry* 13.4, pp. 311–317. ISSN: 1047840X.

Teresi, Jeanne A (2006). "Overview of quantitative measurement methods: Equivalence, invariance, and differential item functioning in health applications". In: *Medical care* 44.11, S39–S49.

The World Bank Group. Tanzania Gender Assessment (2022). Tech. rep. URL: https://documents1.worldbank.org/curated/en/099155103312251069/pdf/P1760510ca289d0400a40e03f6408826f07.pdf.

Thissen D, Steinberg L. Wainer H. (1993). "Detection of differential item functioning using the parameters of item response models. In Differential Item Functioning". In: URL: 10.1093/0195305191.003.0028.

Tong, Eddie M.W. et al. (2010). "Re-examining hope: The roles of agency thinking and pathways thinking". In: *http://dx.doi.org/10.1080/02699930903138865* 24.7, pp. 1207–1215. ISSN: 02699931. DOI: 10.1080/02699930903138865. URL: https://www.tandfonline.com/doi/abs/10.1080/02699930903138865.

UNICEF. Education Budget Brief, (2018). Tech. rep. URL: https://www.unicef.org/tanzania/reports/education-budget-brief-2018.

UNICEF. Education Budget Brief, (2020). Tech. rep. URL: https://www.unicef.org/esa/media/8421/file/UNICEF-Tanzania-Mainland-2020-Education-Budget-Brief-revised.pdf.

UNICEF. Education Fact Sheet:Tanzania, (2017). Tech. rep. URL: https://www.unicef.org/tanzania/media/681/file/UNICEF-Tanzania-2017-Education-fact-sheet.pdf.

Webb, Darren (2007). "Modes of hoping". In: *History of the Human Sciences* 20.3, pp. 65–83. ISSN: 09526951. DOI: 10.1177/0952695107079335.

World Bank, (2021). URL: https://www.worldbank.org/en/country/tanzania/overview#1.

World Faiths Development Dialogue (WFDD). Faith and Development in Focus: Tanzania. Berkely Centre for Religion, Peace and World Affairs, Georgetown University, (2018). Tech. rep. URL: https://berkleycenter.georgetown.edu/publications/faith-and-development-in-focus-tanzania.

World Vision. Breaking the chains of poverty with Empowered World View, (2017). Tech. rep. URL: https://wvusstatic.com/2018/Empowered\%20Worldview.pdf.

World Vision Tanzania, (2022). Tech. rep. URL: https://www.wvi.org/tanzania.

World Vision. Watoto Wetu 2018: World Vision Annual Report, (2018). Tech. rep. URL: https://www.wvi.org/sites/default/files/2019-07/2018\%20ANNUAL\%20REPORT\%20-\%20RESIZED.pdf.

Zanon, Cristian et al. (2016). "An application of item response theory to psychological test development". In: *Psicologia: Reflexao e Critica* 29.1. ISSN: 16787153. DOI: 10.1186/s41155-016-0040-x. URL: http://dx.doi.org/10.1186/s41155-016-0040-x.

Appendix A: Additional chapter. Results on Self-efficacy and Social Cohesion

Our dataset contains questions measuring self-efficacy and social cohesion. Although the primary focus of our study is hope, analyzing the validity of our questions with respect to these two other constructs will allow us to further measure the quality of our survey questions. Specifically, we will test whether the questions in fact measure self-efficacy and cohesion, and whether respondents understand the questions in a similar way. This understanding will be instrumental in the further stages of research.

A.1 Self-Efficacy

Self-efficacy is defined as the belief in one's capacity to accomplish a particular set of actions (Bandura 1982). Self-efficacy is similar to hope in that it is goal-directed and future-oriented. In other words, self-efficacy is a component of the process by which individuals choose objectives and organise their activities to see those goals realised in their lives. Like hope, self-efficacy is cognitive in origin and is centred on one's own well-being and accomplishments. Bandura defines self-efficacy as the notion that one can accomplish something, rather than the idea that one will do something. To think that you can run twenty kilometres is one thing; it is another to find the willpower to actually go out and run those twenty kilometers. The difference is between what one can do versus what one will do. In the meantime hope is characterised by its trait-like characteristics, the original notion of self-efficacy is characterised by its context specificity (Bandura, 1982). While Bandura has expressed concerns about defining self-efficacy as a generic feature, this has not deterred other researchers from

doing so as well. Several measures of general self-efficacy have been created over the course of history (Chen, Gully, and Eden 2001; Jerusalem 1992; Sherer et al. 1982). Using these measures, researchers have discovered that higher levels of general self-efficacy relate to greater psychological well-being. Luszczynska et al. (2005) found that higher general self-efficacy was associated with lower levels of negative affect and depression, as well as higher levels of positive affect and life satisfaction. They found that higher general self-efficacy was associated with higher levels of positive affect and life satisfaction (Luszczynska, Scholz, and Schwarzer 2005).

Self-efficacy is shown to be an important component of behavior change. A self-efficacious person is optimistic about their ability to overcome temptation or starting a new habit. In fact, self-efficacy is a necessary component if these changes are to be seen through. It was found, for example, that self-efficacy improves the odds of successful health behavior change (Bandura et al. 2019; Franko et al. 2008; Neumark-Sztainer et al. 2003). In turn, doubtful individuals fail to either go from intentions to plans or from plans to actions. Self-efficacy helps in achieving our goals because it makes people more likely to plan and take actions. Moreover, self-efficacious people are more confident about trying a novel or difficult behavior and are not set back in the face of obstacles.

The link between hope and self-efficacy has been established in several studies. In the research of Packer et al.(2020), early marriage and childbearing, often motivated by poverty, seemed to be major obstacles to achieving potential objectives for some girls. They witnessed that avoiding marriage and childbearing was closely related to being more hopeful, achieving greater self-efficacy, and being able to take realistic measures to meet future goals. Their results show that respondents who delayed marriage and childbearing were more frequently in school and earning money and were thus typically on track to achieve their goals, and respondents who had high hope and self-efficacy were more frequently delaying marriage and childbearing.

When it comes to content validity and psychometrics, it looks like the Snyder hope scale items cover the same conceptual area as the general self efficacy scale (GSE) (Schwarzer and Jerusalem, 1995). For example, “I know how to handle unforseen problems” from GSE and “I can think of many ways to get out of a jam” from Snyder’s hope scale.

Table A.1 illustrates frequency distribution of self-efficacy items and shows in what order items were asked from respondents. We can see that in each category we have enough observations.

	Self-efficacy items	Strongly disagree	Disagree	Agree	Strongly agree
1	I have the skills and knowledge I need to solve difficult problems	821	1806	2011	746
2	I know how to handle unforeseen problems	834	1873	2030	647
3	I am able to succeed in ways that really matter to me	607	1518	2253	1006
4	I am capable of finding support from others when I need it	200	686	2308	2190
5	I draw inspiration from my spiritual beliefs	153	554	2219	2458
6	I am confident that I can participate in community activities	301	885	2016	2182
7	I am confident that I can contribute to solutions faced by my community	468	1338	1969	1609
8	There are people in my life that I can completely trust	182	672	2170	2360
9	The things I need to solve my problems are readily available to me	1161	2192	1573	458

Table A.1: Frequency distribution for self-efficacy items

Note: 1 - Not at all; 2- Hardly true; 3 - Moderately true; 4- Exactly true	Obs.	Mean	SD	Skew	Kurt	Correlation with total score (Item excluded)
Item 1. I have the skills and knowledge I need to solve difficult problems.	5384	2.50	0.91	-0.05	2.19	0.67
Item 2. I know how to handle unforeseen problems.	5384	2.46	0.89	-0.03	2.24	0.69
Item 3. I am able to succeed in ways that really matter to me.	5384	2.68	0.90	-0.24	2.29	0.66
Item 4. I am capable of finding support from others when I need it.	5384	3.21	0.80	-0.82	3.19	0.60
Item 5. I draw inspiration from my spiritual beliefs.	5384	3.30	0.77	-0.94	3.50	0.56
Item 6. I am confident that I can participate in community activities.	5384	3.13	0.88	-0.74	2.73	0.71
Item 7. I am confident that I can contribute to solutions faced by my community	5384	2.88	0.94	-0.38	2.20	0.71
Item 8. There are people in my life that I can completely trust.	5384	3.25	0.80	-0.87	3.21	0.44
Item 9. The things I need to solve my problems are readily available to me.	5384	2.25	0.89	0.23	2.29	0.45

Table A.2: Descriptive statistics for Self-Efficacy

Table A.2 provides descriptive statistics on self-efficacy questions. 5384 respondents (aged 10–49) participated in the self-efficacy survey. Mean scores of the first three items fall between the categories "hardly true" and "moderately true". Respondents endorsed items from 4 to 6 and 8 more positively by responding "moderately true" or "exactly true". The mean score for item 9 was the lowest (2.25) than any other measures of self-efficacy. Both skewness and kurtosis fall into the acceptable category. The last column shows the correlation between mean item response with the total score of self-efficacy. We see that all items demonstrate a very good correlation. Cronbach's alpha for the self-efficacy survey is 0.79, which falls in the range of good and acceptable levels of reliability.

Figure A.1 illustrates Pearson correlations of the items graphically. As the correlation increases, the color of the boxes changes from red (correlation of -1) to dark blue (correlation of 1). As we can see, no two items are negatively correlated, which means there are no items with opposite meanings. At the same time, we may notice

a strong correlation between items 1 and 2 as well as between items 6 and 7. This suggests that the questions in these pairs ask about strongly related things. For example, items 1 and 2 both ask about personal ability to handle difficult and unforeseen problems. Given that difficult problems are often unforeseen and vice versa, these questions may be asking strongly related, if not same, questions. Likewise, item 6 asks about participation in community activities, and item 7 asks about contributing to solving community problems. In this case solving problems in the community is an important part of participation in community activities, which explains the strong relation between items 6 and 7.

Figure A.1: Pearson correlation of Self-efficacy items

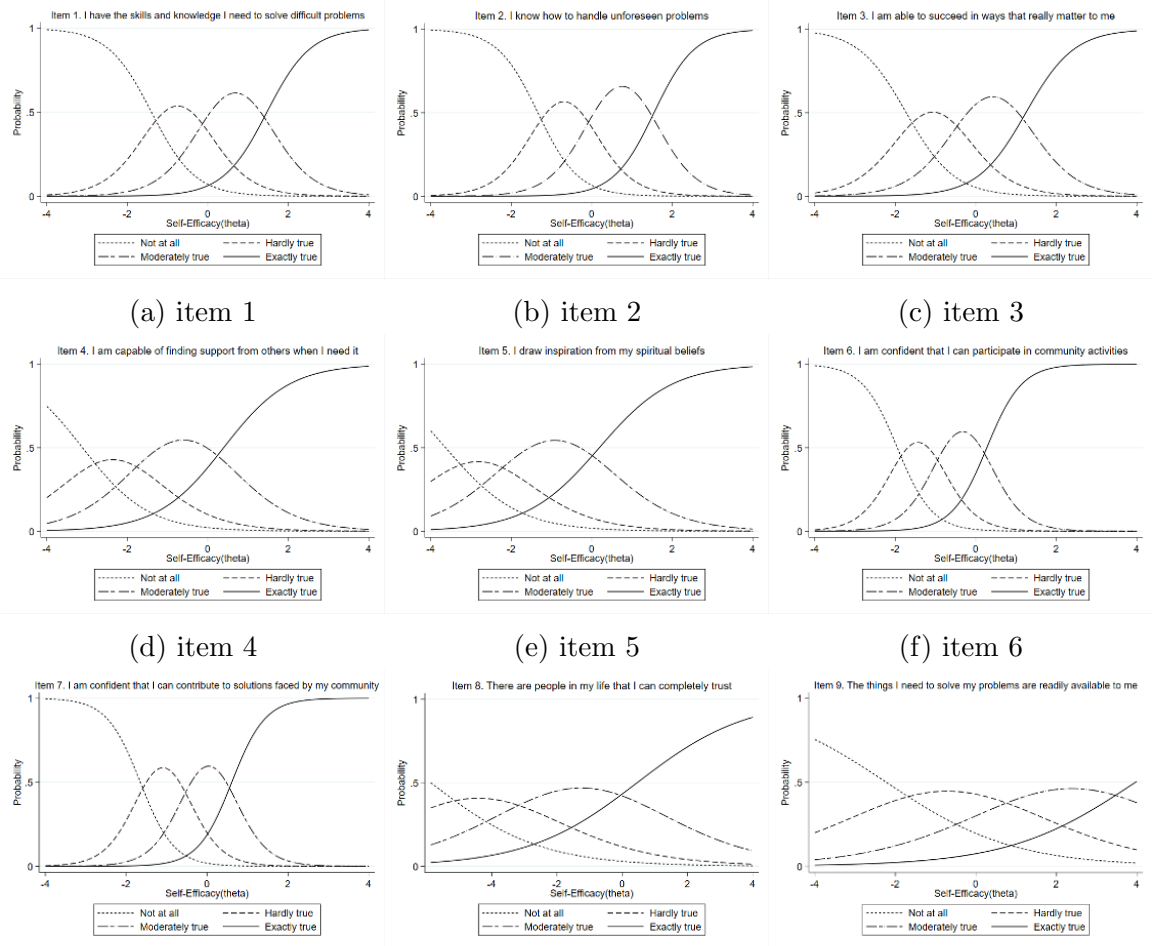
A.1.1 Objective 1: IRT for Self-efficacy

Examining the general self-efficacy measures using IRT in Table A.6, we find that all items can be grouped into four subgroups. The first group consists of Items 6 and 7. We see that Items 6 and 7, related to community, are the most discriminating items, with slope estimates of 2.3 and 2.4, respectively. The next group is Items 1, 2, and 3, which highly differentiate respondents according to their self-efficacy level. Items 4 and 5 are moderately discriminating items, covering respondents with lower or average levels of hope. The last group of two Items, 8 and 9, have a low ability to differentiate between respondents with different levels of self-efficacy. In general, each

self-efficacy item has the most information in the range of modest levels of self-efficacy.

	a	b1	b2	b3
	(SE)	(SE)	(SE)	(SE)
Item 1. I have the skills and knowledge I need to solve difficult problems.	1.84	-1.40	-0.09	1.47
	(0.06)	(0.04)	(0.02)	(0.04)
Item 2. I know how to handle unforeseen problems.	1.98	-1.34	-0.05	1.55
	(0.07)	(0.04)	(0.02)	(0.04)
Item 3. I am able to succeed in ways that really matter to me.	1.67	-1.73	-0.40	1.24
	(0.05)	(0.05)	(0.02)	(0.04)
Item 4. I am capable of finding support from others when I need it.	1.22	-3.11	-1.61	0.39
	(0.04)	(0.11)	(0.05)	(0.03)
Item 5. I draw inspiration from my spiritual beliefs.	1.10	-3.63	-2.02	0.19
	(0.04)	(0.14)	(0.07)	(0.03)
Item 6. I am confident that I can participate in community activities.	2.30	-1.95	-0.92	0.27
	(0.08)	(0.05)	(0.03)	(0.02)
Item 7. I am confident that I can contribute to solutions faced by my community	2.40	-1.65	0.53	0.61
	(0.08)	(0.04)	(0.02)	(0.02)
Item 8. There are people in my life that I can completely trust.	0.59	-5.87	-2.96	0.45
	(0.03)	(0.33)	(0.17)	(0.06)
Item 9. The things I need to solve my problems are readily available to me.	0.63	-2.22	0.81	3.97
	(0.03)	(0.11)	(0.06)	(0.19)

Table A.3: Item GRM Parameter Estimates for Self-efficacy



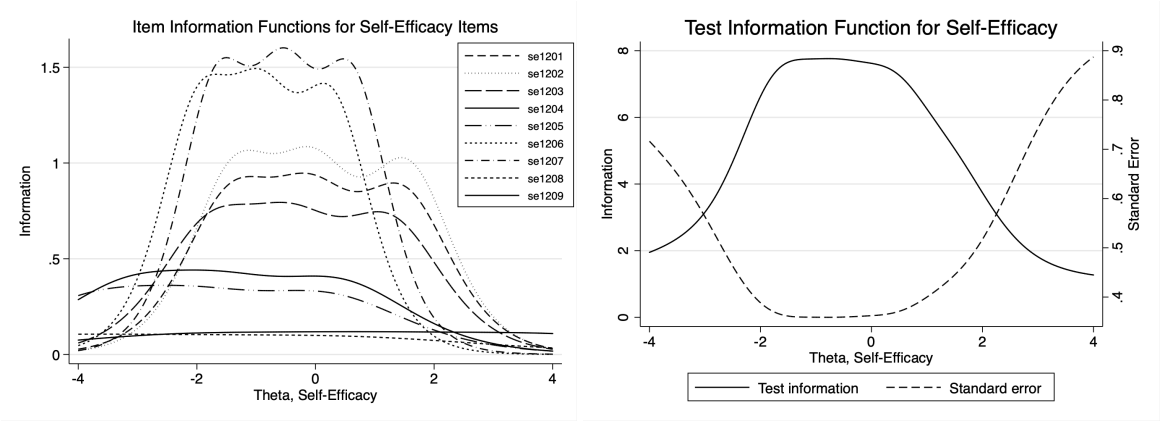
(a) item 1 (b) item 2 (c) item 3

(d) item 4 (e) item 5 (f) item 6

(g) item 7 (h) item 8 (i) item 9

Figure A.2: Category response curves for a four-category graded response model (GRM) item for all 15 items of Social Cohesion. Latent trait (Theta) is shown on the horizontal axis and the probability is shown on the vertical axis.

A.1.2 Information Functions and Total Information Functions



(a) Item Information Function (IIF) (b) Test Information Function (TIF)

Figure A.3: Item Information Function (IIF) and Test Information Function (TIF) of the Self-efficacy. Latent trait is shown on the horizonal axis, and the amount of information for each item for IIF and the amount of information for all items and standard error yielded by the test at any trait level are shown on the vertical axis.

Looking at item information functions in Figure A.3a, we can say Items 1 to 7 provide good information across people with different levels of self-efficacy, especially at lower and modest levels. Items 8 and 9 provide the least information and are suggested for revision. The graph of test information function in Figure A.3b tells us that the self-efficacy items are most informative within two standard deviations from the mean.

A.1.3 Objective 2: DIF for Self-efficacy

	Gender (SE) (1-male; 2-female)	Region (SE) (0-Kigoma; 1-Iringa)	Age (SE) (0-adoles;1-adult;)	Food Security (SE) (0-secure, 1-no)	Access to protected water (SE) (0-yes, 1-no)	Treatment (SE) (0-control; 1-intervention)
Item 1. I have the skills and knowledge I need to solve difficult problems.	-0.02 (0.06)	-0.12 (0.07)*	-0.03 (0.07)	0.04 (0.08)	-0.24 (0.06)***	0.14 (0.06)**
Item 2. I know how to handle unforeseen problems.	-0.15 (0.06)*	0.02 (0.07)	0.32 (0.07)***	-0.03 (0.08)	-0.31 (0.07)***	0.14 (0.06)**
Item 3. I am able to succeed in ways that really matter to me.	-0.05 (0.06)	0.30 (0.06)***	0.11 (0.06)*	-0.33 (0.07)***	-0.18 (0.06)**	0.06 (0.06)
Item 4. I am capable of finding support from others when I need it.	0.04 (0.06)	0.38 (0.06)***	-0.27 (0.07)***	0.33 (0.07)***	0.21 (0.06)**	-0.08 (0.06)
Item 5. I draw inspiration from my spiritual beliefs.	0.25 (0.06)***	0.45 (0.06)***	-0.27 (0.07)***	-0.19 (0.07)**	0.32 (0.06)***	-0.09 (0.06)
Item 6. I am confident that I can participate in community activities.	-0.09 (0.06)	0.23 (0.07)**	1.03 (0.07)***	0.19 (0.08)*	0.36 (0.07)***	0.04 (0.06)
Item 7. I am confident that I can contribute to solutions faced by my community	-0.39 (0.06)***	-0.14 (0.07)*	1.47 (0.07)***	-0.22 (0.08)**	0.04 (0.07)	-0.05 (0.06)
Item 8. There are people in my life that I can completely trust.	-0.09 (0.06)*	-0.10 (0.06)	-0.84 (0.06)	-0.03 (0.07)	0.20 (0.06)**	0.02 (0.05)
Item 9. The things I need to solve my problems are readily available to me.	0.09 (0.05)*	0.41 (0.06)***	-0.67 (0.06)***	-0.31 (0.07)	0.04 (0.06)	0.03 (0.05)

*Note: *=p<0.1; ** p<0.05 ;***= p<0.01*
Food security: In the past 12 months how did your household meet daily food needs?
Water access: What is the main water source for your household?

Table A.4: DIF for Self-efficacy

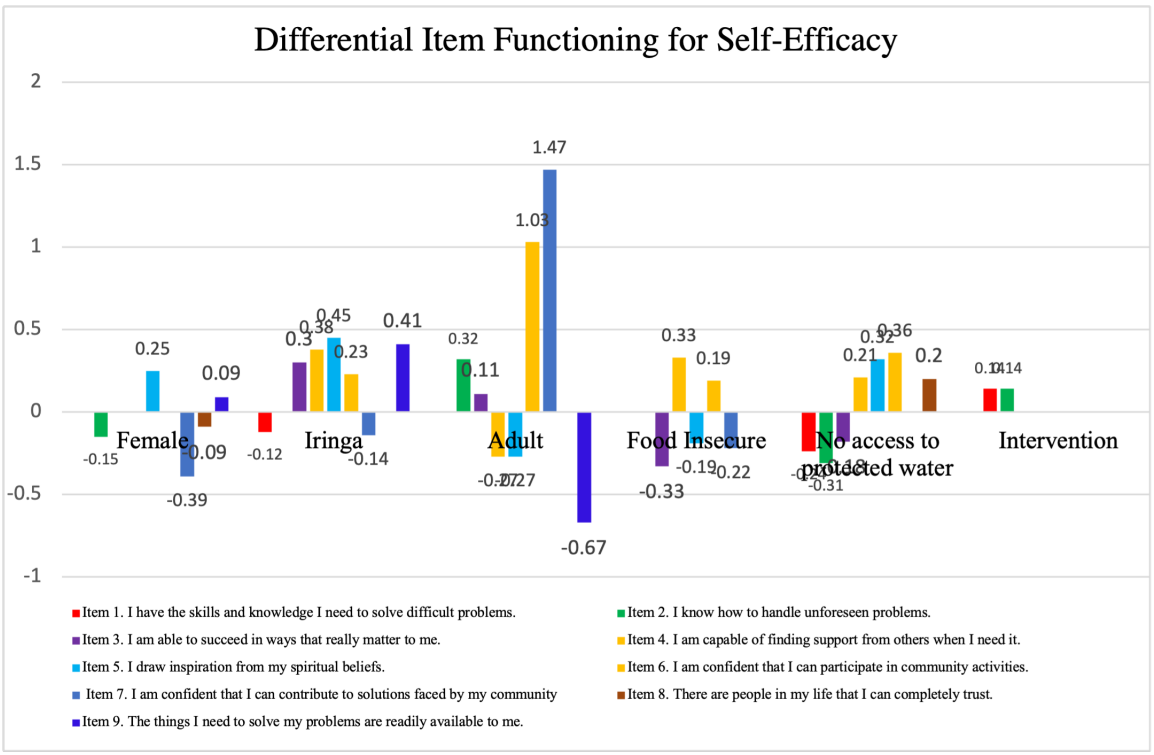


Figure A.4: DIF for Self-efficacy items

Table A.4 provides DIF figures across gender, region, age category, food security, access to water, and belonging to the intervention group and figure A.4 visually presents them for convenience. Observing DIF across genders, we find that females are less likely to endorse items 2 and 7 than males, meaning that for the same level of self-efficacy, women are less likely to believe they can handle unforeseen problems and less confident that they can contribute to solutions faced by their communities. At the same time, we find that women are more likely to draw inspiration from their spiritual beliefs. It could be that women can be as self-efficacious as men, relying on their spiritual beliefs while not having the same opportunities to personally affect their communities' outcomes. This agrees with the idea in the DIF for hope that women are more likely to derive empowerment from external forces, while men rely on their internal ability to generate hope and self-efficacy. Furthermore, given the same level of self-efficacy, respondents from Iringa have a higher chance of endorsing items 3-6 than their Kigoma counterparts. Thus, for the same level of self-efficacy, Iringa respondents are more likely than Kigoma residents to feel they can succeed in meaningful ways, find support, draw inspiration from spiritual beliefs, and participate

in community activities. In line with the DIF result for hope, this may suggest that it takes higher levels of these traits to be as self-efficacious in Iringa as in Kigoma, meaning that self-efficacy in Iringa requires more individual ability. Notably, given the same level of self-efficacy as Kigomans, Iringans had a lower belief in their ability to contribute to solving the community's challenges, which may suggest that the concept of self-efficacy in Iringa is less community oriented. We can also see that adults are more likely to endorse items about community activities. They are more likely to agree that they feel confident in participating in community activities and contributing to solutions faced by them than adolescents. Adults are also less certain of being able to find support from their community and spirituality and less likely to believe that they are adequately equipped to solve their problems. The suggestion is that adults rely more on themselves than on others' help, as opposed to adolescents, given the same level of self-efficacy. Respondents who are food insecure and have no access to protected water are more likely to agree that they can find support from others and feel confident participating in community activities. Unfortunately, food insecure individuals also feel less empowered in other important ways. Thus, they are less likely to believe they can succeed in ways important to them (Item 3), less likely to say they draw inspiration from their spiritual beliefs (Item 5), and feel empowered to contribute to the community (Item 7). The absence of DIF for the intervention group (compared to the control group) indicates that the probability of a person's response to items depends only on her or his hope level and not on her or his belonging to the focal group.

A.2 Social Cohesion

Hope can be affected by various domains of our lives, such as family members, close friends, the working environment, the society where we live, the climate and others. Moreover, assumptions and expectations about what can and can't be achieved are heavily influenced by social norms and beliefs (Benzein and Saveman 1998). We can learn to be hopeful, especially in childhood (Snyder, 2000a; Webb, 2013). Hope can also come from having trust in others, which trust gives us support, safety, and confidence (Stevenson and Peterson, 2016). It can also be important to have well-

functioning social institutions like judicial systems or police forces in order to grow and keep hope. These institutions provide the safety needed to live peacefully and make progress both for themselves and for their communities (Drahos 2004; Stevenson and Peterson 2016) (Drahos, 2004; Stevenson Peterson, 2016). Also, it is a political issue because in society, hope might not always be given equally, with some groups, such as minorities or underprivileged groups, being granted less hope than others. So, hope can become a public, social, or shared hope in reaction to societal or political circumstances (Bar-Tal 2001; Kleist and Jansen 2016a).

In the mindset survey, we have a set of social cohesion questions that have added to our understanding of hope. Previous works reveal a disparity in perceptions of what constitutes social cohesion. Social cohesion is described as "the vertical and horizontal interaction among members of society as defined by a set of attitudes and norms that include trust, a sense of belonging, and a willingness to participate and help" (Chan, To, and Chan 2006). It is characterised by the lack of conflict or crime (Klein 2014), the strength of connection (Braaten 1991), and a certain degree of stability (Talcott 2013). Numerous definitions of economic parts of society have been adopted internationally, including general well-being and equal representation/opportunities in society (Jeannotte 2003). In these definitions, things like the well-being of group members, shared values like trust, and equal chances in society are taken into account.

In this work, we rely on Langer et al. (2017) understanding of social cohesion. They crystallised the definitions of social cohesion as highlighting the following: a fair distribution of power and material resources; links between individuals; shared cultures and morals; and bonds of reciprocity to keep people together. Social cohesion can also be on any scale – interethnic, international, between and across groups, or between individuals and institutions. Groups can be classified according to ethnic origin, religion, race, and native or immigrant status. Langer et al. (2017) elaborate that social cohesion consists of a triangle of trust, inequality, and identity.

The social cohesion survey we use is a 15-item Likert-type scale ranging from 1, "strongly disagree", to 4, "strongly agree". It consists of 11 positive and 4 negative items. The 4 negative items are recorded by giving 1 to "strongly agree" and 4 to "strongly disagree". 5201 respondents (aged 10–49) answered to all questions of social cohesion survey. Table A.5 illustrates frequency distribution of social cohesion items

and shows in what order items were asked from respondents. We can see that in each category we have enough observations.

	Social Cohesion items	Strongly disagree	Disagree	Agree	Strongly agree
1	People in this community readily help each other	1145	3043	767	404
2	People in this community tend to trust each another	908	3137	979	335
3	People in this community actively care for people outside of their family who are poor, weak or vulnerable	757	2403	1386	806
4	People in this community work together to improve the well-being of all children	571	2576	1441	753
5	The activities that are being implemented in this community to improve the well-being of children are the most important	1162	3147	842	183
6	People in this community wait for government or NGO support in times of need	480	1973	1876	974
7	Women take leadership roles in our community councils, or in community groups	1088	3167	807	264
8	Our leaders listen to input from everyone in the community, including the most vulnerable groups, when making a decision	1055	3403	657	188
9	Our leaders work hard to improving the well-being of children in this community	868	3267	975	200
10	Our leaders are able to obtain assistance from outside the community to improve the lives of our children	989	3397	761	156
11	In this community, when children are malnourished, we wait for the government to solve the problem	196	1271	2262	1574
12	In this community, if a school building needs some repairs, we wait for other outsiders to come and fix it	362	1590	2169	1205
13	In this community, when children drop out of school, we are not able to do anything about it	392	1447	2287	1199
14	When conflicts arise among the members of this community, we are usually able to solve the problem	1722	3047	396	115
15	People in my community have close friends of different religions	1669	2877	626	170

Table A.5: Frequency distribution for social cohesion items

Table A.6 shows that the mean scores of items ranged from 2.55 to 3.21 which is close to the category "agree" on majority of positive items. Mean scores of items 3

and 4 were the lowest (2.58 and 2.55). Mean score for negative items ranged from 2.63 to 2.99. All items except Item 12 are slightly left skewed and all are in the acceptable range. The last column on item correlation with the total score shows that all items except item 15 have a good (Items 10, 12, 13 and 14) or a very good discrimination.

Positive items *Note: 1 - Strongly Disagree; 2- Disagree; 3 - Agree; 4- Strongly Agree*	Obs.	Mean	SD	Skew	Kurt	Correlation with total score (Item excluded)
Item 1. People in this community readily help each other.	5201	2.92	0.81	-0.71	3.32	0.58
Item 2. People in this community tend to trust each another.	5201	2.86	0.77	-0.61	3.33	0.57
Item 3. People in this community actively care for people outside of their family who are poor, weak or vulnerable.	5201	2.58	0.91	-0.27	2.27	0.52
Item 4. People in this community work together to improve the well-being of all children.	5201	2.55	0.87	-0.32	2.41	0.56
Item 5. The activities that are being implemented in this community to improve the well-being of children are the most important.	5201	2.99	0.72	-0.54	3.45	0.60
Item 6. Women take leadership roles in our community councils, or in community groups.	5201	2.96	0.74	-0.65	3.58	0.50
Item 7. Our leaders listen to input from everyone in the community, including the most vulnerable groups, when making a decision.	5201	3.01	0.68	-0.67	4.12	0.61
Item 8. Our leaders work hard to improving the well-being of children in this community	5201	2.91	0.70	-0.52	3.55	0.65
Item 9. Our leaders are able to obtain assistance from outside the community to improve the lives of our children.	5201	2.99	0.67	-0.56	3.91	0.64
Item 10. When conflicts arise among the members of this community, we are usually able to solve the problem.	5201	3.21	0.67	-0.71	4.10	0.38
Item 11. People in my community have close friends of different religions.	5201	3.13	0.74	-0.70	3.50	0.41
Negative statements *Note: 1 - Strongly Agree; 2- Agree; 3 - Disagree; 4- Strongly Disagree*						
Item 12. People in this community wait for government or NGO support in times of need.	5201	2.63	0.89	0.00	2.21	0.30
Item 13. In this community, when children are malnourished, we wait for the government to solve the problem.	5201	2.99	0.83	-0.36	2.37	0.27
Item 14. In this community, if a school building needs some repairs, we wait for other outsiders to come and fix it.	5201	2.80	0.87	-0.22	2.29	0.24
Item 15. In this community, when children drop out of school, we are not able to do anything about it.	5201	2.81	0.87	-0.29	2.39	0.08

Table A.6: Social cohesion, community ownership and trust (reverse coded for negative items)

We see from Figure A.5 that almost all positive items are positively correlated between each other, correlations ranging between 0.15 to 0.50 which is within the acceptable range (Briggs 1986). Correlation values larger than 0.50 are also found

between Items 1 and 2, and between Items 7 and 8. This indicates that the items are repetitious. Item 1 is asking about the level of help in the community and Item 2 is about trust in the community, which are fairly related concepts. Likewise, Items 7 and 8 are statements about community leaders working for the community and helping the community, which are very closely related ideas. Negative statements are independent from or negatively correlated with positive statements. Item 6 asks whether women take leadership roles in the community. This item correlates with leaders working hard and listening to inputs from the community. Item 10 stating "When conflicts arise among the members of this community, we are usually able to solve the problem", does not correlate well with other items except Item 11, "People in my community have close friends of different religions". This is interesting as it may suggest that communities with close ties between people of different religions may be better equipped to solve community conflicts.

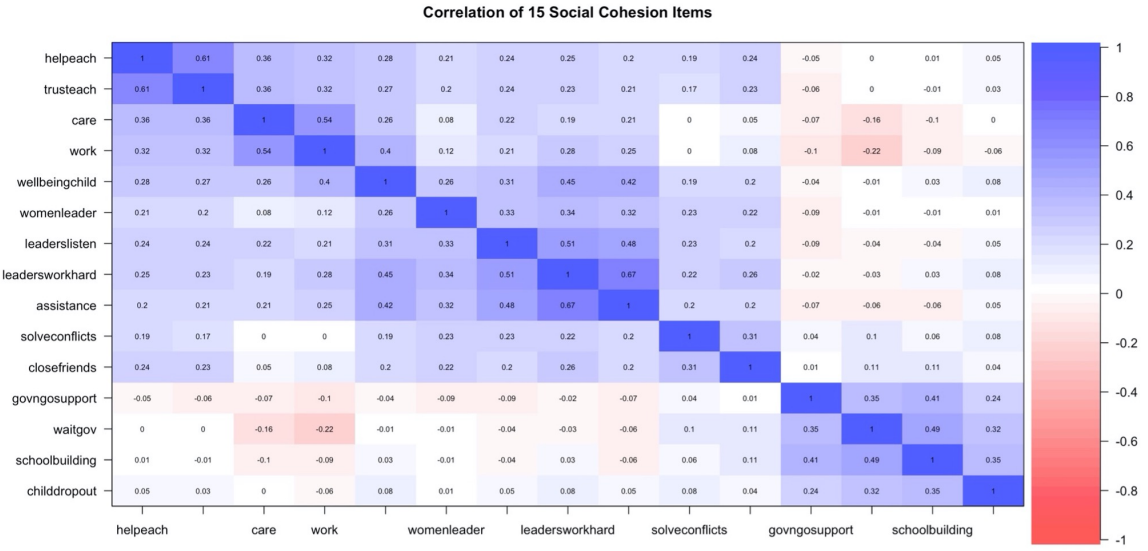


Figure A.5: Pearson correlation for Social Cohesion

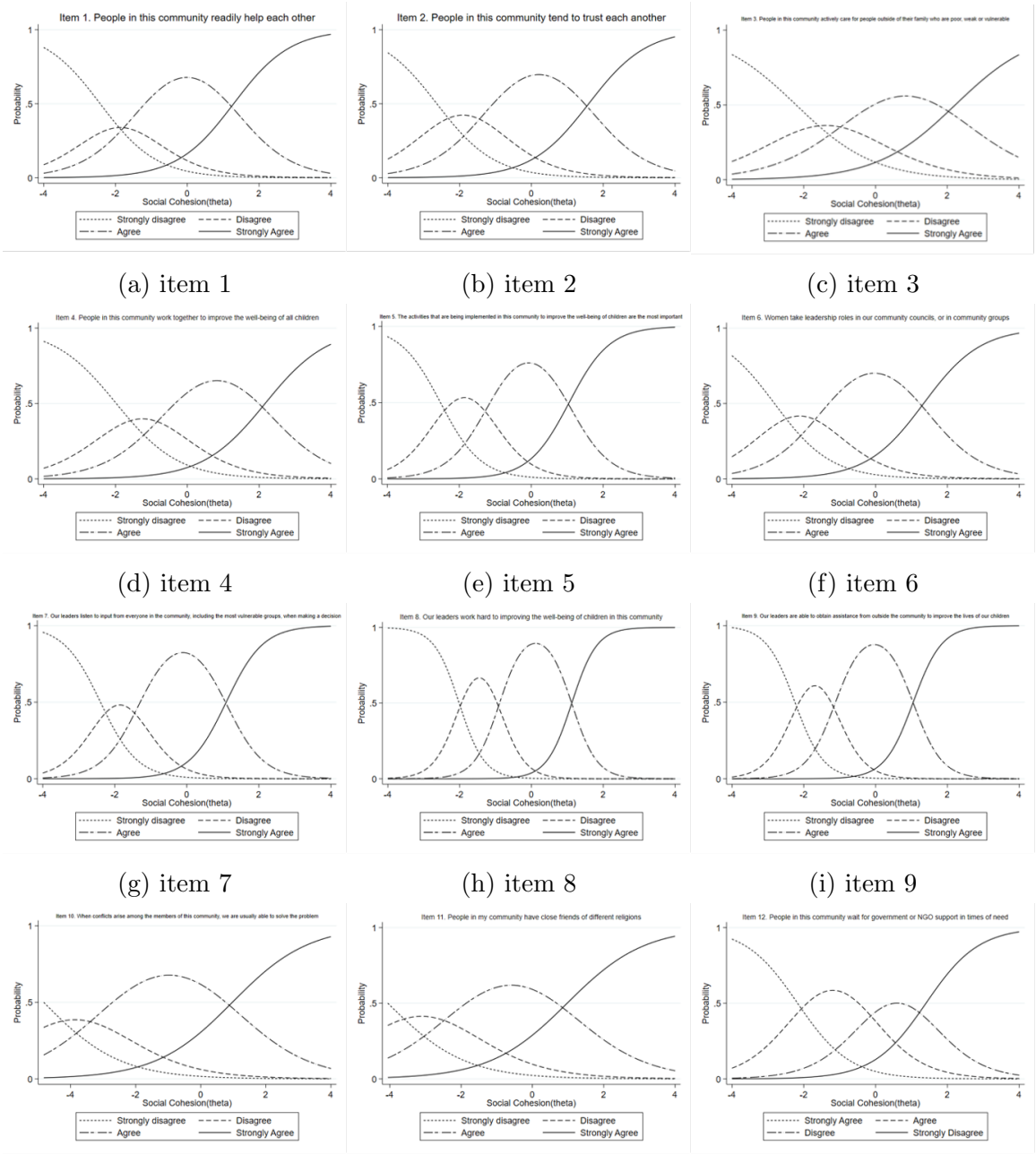
A.2.1 Objective 1: IRT for Social Cohesion

An internal consistency estimate is calculated for all 15 items, which include positive and negative statements, using Cronbach's Alpha with $\alpha = 0.73$. The estimated parameters of IRT show that the value of slope parameters ranged from 0.85 to 2.86,

which is moderate to very high. Items 1, 2, 3, 4, 6, 10, 11 and 15 have moderate discrimination. The location estimates of the items shown in table A.7 range from -4.82 to 2.22, indicating a good variation in item location.

	a	*b1*	*b2*	*b3*
	(SE)	*(SE)*	*(SE)*	*(SE)*
Item 1. People in this community readily help each other.	1.27	-2.43	-1.31	1.29
	(0.04)	(0.08)	(0.04)	(0.05)
Item 2. People in this community tend to trust each another.	1.24	-2.64	-1.18	1.59
	(0.04)	(0.08)	(0.04)	(0.05)
Item 3. People in this community actively care for people outside of their family who are poor, weak or vulnerable.	0.91	-2.21	-0.55	2.22
	(0.04)	(0.08)	(0.04)	(0.09)
Item 4. People in this community work together to improve the well-being of all children.	1.16	-1.98	-0.52	2.17
	(0.04)	(0.06)	(0.03)	(0.07)
Item 5. The activities that are being implemented in this community to improve the well-being of children are the most important.	1.78	-2.52	-1.19	1.05
	(0.05)	(0.07)	(0.03)	(0.03)
Item 6. Women take leadership roles in our community councils, or in community groups.	1.26	-2.82	-1.41	1.34
	(0.04)	(0.09)	(0.05)	(0.05)
Item 7. Our leaders listen to input from everyone in the community, including the most vulnerable groups, when making a decision.	1.94	-2.39	-1.32	1.09
	(0.06)	(0.06)	(0.03)	(0.03)
Item 8. Our leaders work hard to improving the well-being of children in this community	2.86	-2.01	-0.89	1.12
	(0.09)	(0.04)	(0.02)	(0.03)
Item 9. Our leaders are able to obtain assistance from outside the community to improve the lives of our children.	2.51	-2.25	-1.12	1.05
	(0.08)	(0.05)	(0.03)	(0.03)
Item 10. When conflicts arise among the members of this community, we are usually able to solve the problem.	0.85	-4.82	-2.91	0.97
	(0.04)	(0.22)	(0.12)	(0.05)
Item 11. People in my community have close friends of different religions.	0.93	-4.06	-2.15	0.97
	(0.04)	(0.16)	(0.08)	(0.05)
Negative statements				
Item 12. People in this community wait for government or NGO support in times of need.	1.37	-2.17	-0.21	1.4
	(0.05)	(0.06)	(0.03)	(0.05)
Item 13. In this community, when children are malnourished, we wait for the government to solve the problem.	2.02	-2.36	-0.77	0.68
	(0.07)	(0.06)	(0.03)	(0.03)
Item 14. In this community, if a school building needs some repairs, we wait for other outsiders to come and fix it.	2.56	-1.79	-0.44	0.88
	(0.1)	(0.04)	(0.02)	(0.03)
Item 15. In this community, when children drop out of school, we are not able to do anything about it.	1.23	-2.53	-0.72	1.24
	(0.04)	(0.08)	(0.03)	(0.05)

Table A.7: Item GRM Parameter Estimates for Social Cohesion



(a) item 1 (b) item 2 (c) item 3

(d) item 4 (e) item 5 (f) item 6

(g) item 7 (h) item 8 (i) item 9

(j) item 10 (k) item 11 (l) item 12

Figure A.6: Category response curves for a four-category graded response model (GRM) item for all 15 items of Social Cohesion. Latent trait (Theta) is shown on the horizontal axis and the probability is shown on the vertical axis.

A.2.2 Information Functions and Total Information Functions

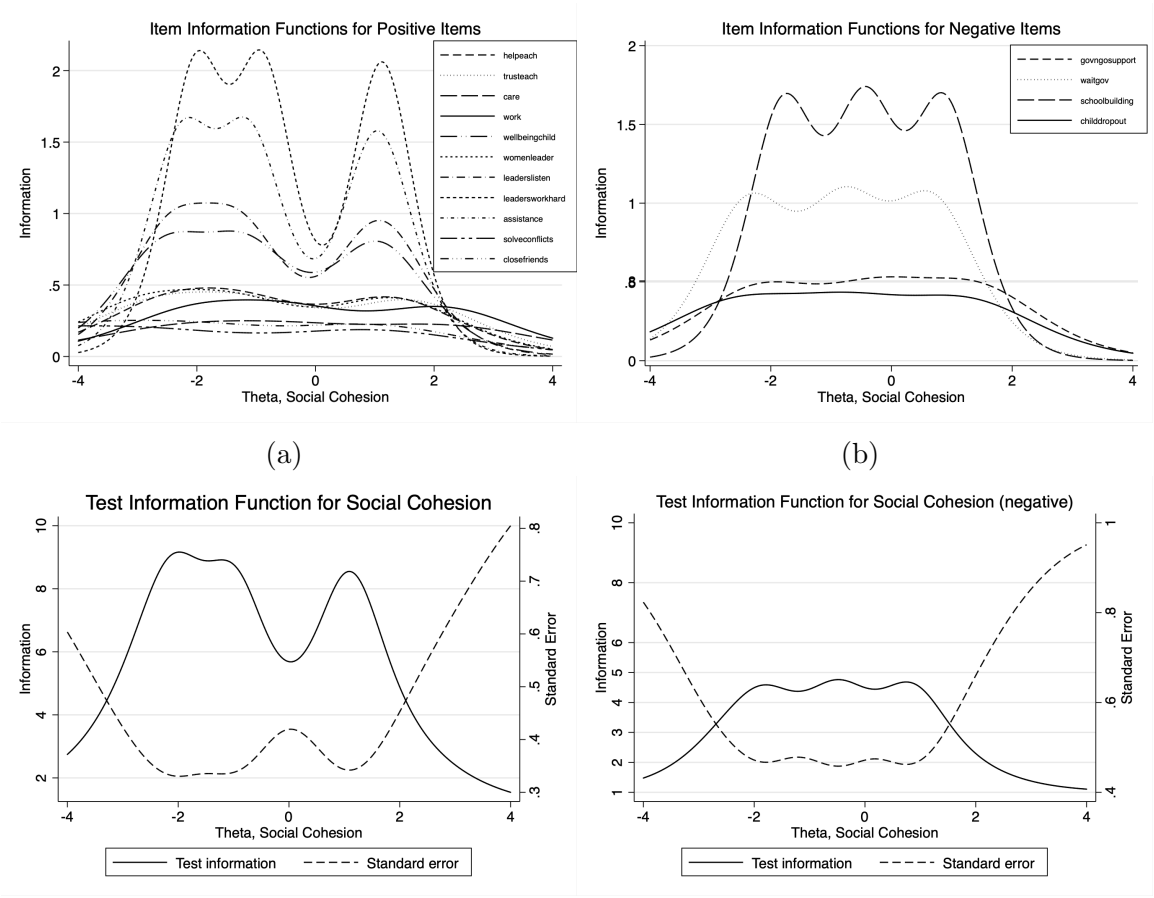


Figure A.7: Category response curves for a four-category graded response model (GRM) item for all 15 items of Social Cohesion. Latent trait (Theta) is shown on the horizontal axis and the probability is shown on the vertical axis.

IFs and TIF show a good coverage of a wide range of the latent social cohesion level among respondents (see Figure A.7). The items are most informative around a theta level of -2. There is an additional peak at values around theta equal to +1.5. We do not observe as much variation in informativeness for positive and negative items in the social cohesion survey as we did with hope. Items 3, 10, and 11 overlap and contribute the least amount of information, while Items 8, 9 and 14 provide most of the information. Interestingly, items are least informative around the social cohesion level of zero. The source of this drop in informativeness is unclear at this point.

A.2.3 Objective 2: DIF for Social Cohesion

	Gender (SE)	*Region (SE)*	*Age (SE)*	*Food Security (SE)*	*Access to protected water (SE)*	*Treatment (SE)*
Item 1. People in this community readily help each other.	-0.02 (0.06)	-0.2 (0.06)**	-0.19 (0.7)	0.1 (0.08)	0.08 (0.06)	0.16 (0.06)**
Item 2. People in this community tend to trust each another.	-0.06 (0.06)	0.04 (0.06)	-0.31 (0.06)***	0.01 (0.08)	0.12 (0.06)*	0.08 (0.06)
Item 3. People in this community actively care for people outside of their family who are poor, weak or vulnerable.	0.03 (0.05)	0.01 (0.06)	-0.56 (0.06)***	-0.94 (0.07)***	-0.08 (0.06)	-0.03 (0.05)
Item 4. People in this community work together to improve the well-being of all children.	-0.02 (0.06)	0.13 (0.06)**	-0.57 (0.06)***	-0.53 (0.07)***	-0.1 (0.06)*	0.00 (0.05)
Item 5. The activities that are being implemented in this community to improve the well-being of children are the most important.	-0.04 (0.06)	0.09 (0.07)	0.22 (0.07)**	-0.26 (0.08)**	0.02 (0.07)	0.07 (0.06)
Item 6. Women take leadership roles in our community councils, or in community groups.	0.07 (0.06)	0.33 (0.07)***	0.28 (0.06)***	0.11 (0.08)	0.12 (0.06)	-0.18 (0.06)**
Item 7. Our leaders listen to input from everyone in the community, including the most vulnerable groups, when making a decision.	-0.06 (0.07)	0.01 (0.07)	0.1 (0.07)	0.30 (0.09)***	-0.20 (0.07)***	-0.02 (0.07)
Item 8. Our leaders work hard to improving the well-being of children in this community	0.05 (0.08)	0.18 (0.08)**	0.03 (0.08)	-0.05 (0.10)	-0.03 (0.08)	-0.03 (0.08)
Item 9. Our leaders are able to obtain assistance from outside the community to improve the lives of our children.	0.01 (0.07)	-0.05 (0.08)	-0.15 (0.08)**	-0.08 (0.09)	-0.23 (0.08)***	0.04 (0.07)
Item 10. When conflicts arise among the members of this community, we are usually able to solve the problem.	-0.12 (0.06)*	-0.19 (0.07)***	0.64 (0.06)***	0.72 (0.08)***	0.25 (0.06)***	-0.12 (0.06)**
Item 11. People in my community have close friends of different religions.	-0.13 (0.06)**	0.31 (0.06)***	0.35 (0.06)***	0.13 (0.07)*	0.25 (0.06)***	-0.05 (0.05)
Negative statements						
Item 12. People in this community wait for government or NGO support in times of need.	0.07 (0.05)	-0.18 (0.06)**	0.43 (0.05)***	-0.24 (0.07)***	-0.05 (0.06)	0.13 (0.05)***
Item 13. In this community, when children are malnourished, we wait for the government to solve the problem.	0.02 (0.05)	-0.20 (0.06)***	0.51 (0.05)***	0.27 (0.07)***	0.16 (0.06)***	0.13 (0.05)***
Item 14. In this community, if a school building needs some repairs, we wait for other outsiders to come and fix it.	0.00 (0.05)	-0.15 (0.06)***	0.70 (0.05)***	-0.19 (0.07)***	0.22 (0.06)***	0.10 (0.05)*
Item 15. In this community, when children drop out of school P we are not able to do anything about it.	0.08 (0.05)	-0.54 (0.06)***	0.27 (0.05)***	-0.61 (0.07)***	0.13 (0.06)**	-0.06 (0.05)

Note: *=p<0.1; ** p<0.05 ;***= p<0.01
Food security: In the past 12 months how did your household meet daily food needs?
Water access: What is the main water source for your household?

Table A.8: DIF for gender, age, region, food security, access to water and treatment for Social Cohesion

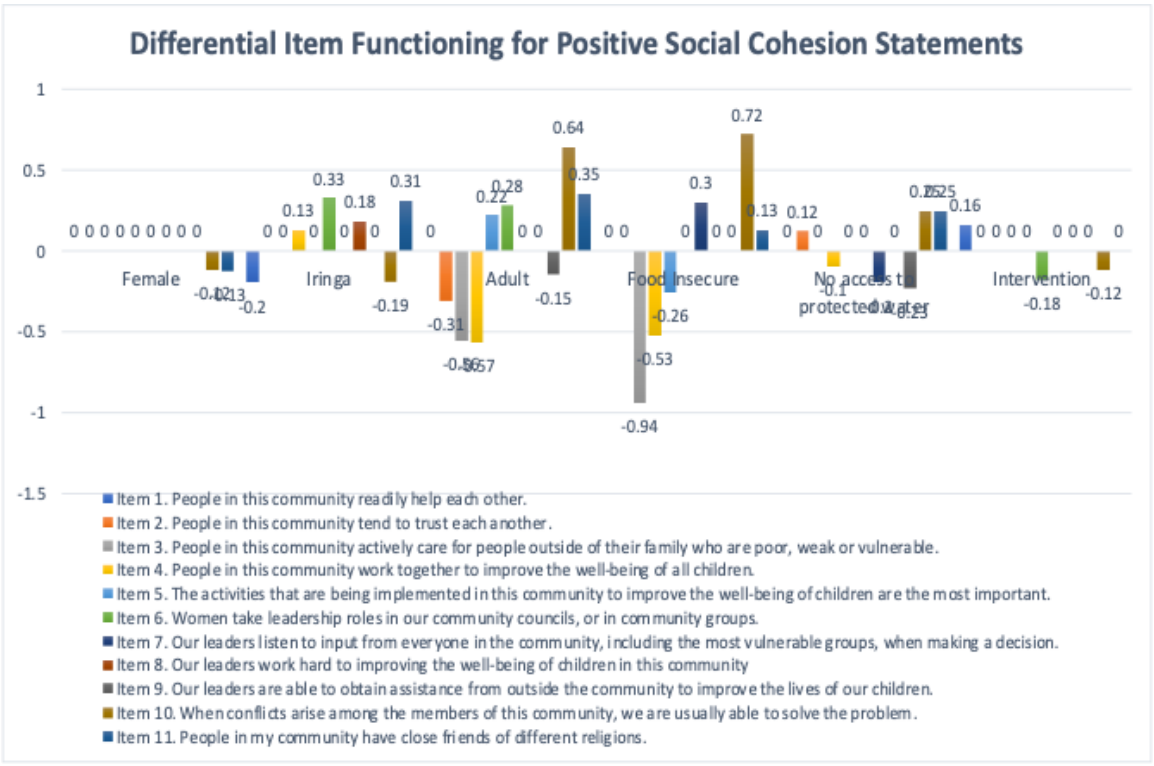


Figure A.8: DIF for positive social cohesion items across categories

Table A.8 reports DIF for gender, age, region, food security, access to water and treatment for Social Cohesion and Figure A.8 illustrates significant DIFs for positively stated items across the categories of gender, region, age, food security, access to protected water, and belonging to the intervention group. The first thing we notice by looking at Figure A.8 is that there are almost no DIFs for any items based on gender, suggesting that male and female respondents understand the questions regarding social cohesion in a similar way. Similarly, there is no DIF for Items 1, 2, 6, 8, and 9 among people who could meet their daily food needs and those who could not, and between those who have no access to protected water and those who have. Also, there is no DIF for any items based on being in the intervention group except for items 6 and 10. Given the same level of social cohesion, people in the treatment group are less likely to agree that women can be leaders in the community and that people in the community can solve their own problems. However, there are DIFs across the remaining categories, suggesting that the social cohesion items may not be invariant. We can see that respondents from Iringa with the same level of social cohesion are more likely to agree with Items 4, 6, 8, and 11 but less likely to agree with Item 10.

Adults are less likely to agree that there is trust in the community, that people take care of poor and vulnerable people, and that they work together to improve the well-being of all children. Item 3, which is "taking care of the poor, weak, or vulnerable," exhibited significant DIF for respondents who are food insecure. They are much less likely to agree that people in the community take care of people outside of their family who are poor, weak, or vulnerable. The same is true with children-related Items 3 and 4. But they are more likely to agree that they usually solve the problems that arise among members of their community themselves. The presence of DIFs indicates that social cohesion items may need to be reviewed before their future use.

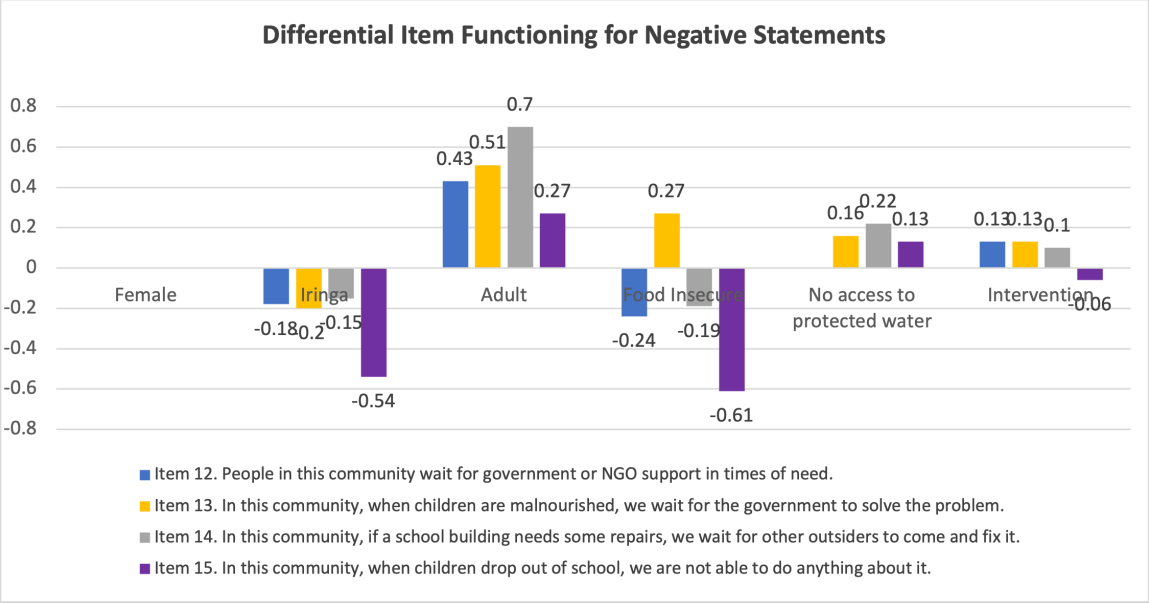


Figure A.9: DIF for negative social cohesion items across categories

Figure A.9 illustrates negative DIF for cohesion items by the categories. The first thing to notice is that there is no DIF for gender. Thus, social cohesion items are understood the same way by both males and females. However, there are DIFs present across other categories. People from Iringa are less likely to disagree with all negative statements compared to Kigoma, given the same level of social cohesion. Furthermore, adults are more likely to disagree with negative statements, especially in terms of item 14. This means that adults feel more optimistic about the level of social cohesion in their community compared to adolescents. Furthermore, food insecure individuals are less likely to disagree with items 12, 14 and 15, while respondents with no access

to protected water tended to disagree with items 13, 14 and 15. Finally, there is a very small DIF between the treatment and control groups. Respondents from the treatment group are more likely to disagree with negative 12-14 items.

A.2.4 Conclusion

DIF analysis for self-efficacy items reveals the presence of DIFs across gender, region, age category, food security, and access to water. With regards to gender, we find that given the same level of self-efficacy, women are more likely to derive empowerment from external forces, while men rely on their internal ability to generate hope and self-efficacy. This could mean that men and women have a different understanding of self-efficacy. As for the regional differences, Iringa respondents are more likely than Kigoma residents to feel they can succeed in meaningful ways, find support, draw inspiration from spiritual beliefs, and participate in community activities, given the same level of self-efficacy. This may suggest that it takes higher levels of these traits to be as self-efficacious in Iringa as in Kigoma, meaning that self-efficacy in Iringa requires more individual ability. Across age, groups, adults are more like to be self-reliant and community-oriented than adolescents, given the same level of self-efficacy. There is also a DIF for food security status. Food insecure individuals feel less empowered to succeed and contribute to their community. Finally, we find no DIF in the category of belonging to the treatment group. This tells us that the division of the sample into the control and treatment groups is arbitrary, and the groups are statistically not very different. The results for DIF analysis for social cohesion are similar to those for self-efficacy except that there is no DIF for gender. This implies that male and female respondents understand the questions regarding social cohesion in a similar way. However, there are DIFs in the remaining categories, suggesting that the social cohesion items may not be invariant. Across regions, we see positive DIFs for Iringa. Particularly we find that given the same level of social cohesion, Iringa respondents indicated that women take leadership roles in the community and people have close friends from other religions. Furthermore, we find that adults and food insecure individuals tend to be more pessimistic about the state of social cohesion in their community. As with self-efficacy, the presence of DIFs indicates that social cohesion items may need to be reviewed before proceeding with the study.

Appendix B: Multidimensionality

The objective of this chapter is to apply different IRT models to explore the dimensionality of the latent trait of hope in our survey. The use of IRT requires the assumption of unidimensionality to hold. When a survey is unidimensional, then the questions measure one latent trait and no other knowledge is required to answer those questions correctly. For example, a unidimensional test of mathematical ability should contain only math questions. If the unidimensionality assumption is broken, then the latent trait is measured by a question that measures something different, for example, by a literature question. This literature question will be interpreted as a 'difficult' math question by IRT if only a few people answer it correctly; and those who answer it correctly will be scored as having high mathematical ability, substantially skewing the results of the survey. To avoid these kinds of problems, tests of dimensionality recently started to be introduced. Yet, it is still common practise in the hope literature to use unidimensional models to analyze the data. In this work, we conduct tests of dimensionality for our hope survey. First, we conduct exploratory analysis consisting of principal component and exploratory factor analysis to figure out how many dimensions of hope we may have. Second, we estimate omega, which helps us to check the internal consistency and perform exploratory factor analysis. Finally, after identifying the factors of the survey, we will build unidimensional and multidimensional models and compare them for the best data fit.

To investigate whether multidimensionality exists in the hope survey, the exploratory approaches of principal component analysis (PCA) and exploratory factor analysis (EFA) were applied first. PCA reduces the number of observed variables to a smaller number of principal components that explain the majority of the observed variables' variation. EFA constructs a linear model of a collection of variables by finding the underlying latent factors. We will also calculate omega, which is a re-

liability test to measure internal consistency and which also performs EFA. When multidimensionality may be present, it is advised that omega should be used (Chen et al. 2012); and several researchers advocate for the increased use of omega over alpha (Dunn, Baguley, and Brunsden 2014; Schweizer 2011).

Confirmatory factor analysis (CFA) is used when we already know about the underlying structure of the hope trait. We may proceed with particular IRT models for confirmatory analyses after we have an understanding of the probable latent hope dimension(s) of data from the EFA.

B.0.1 Results

The first step was to perform PCA. One method for calculating the number of components is to choose those with Eigenvalues larger than 1. This score indicates that these factors explain more than the mean of the total variance in the items (Guttman, 1954; Kaiser, 1960).

Component	Eigenvalue	Variance	Explained
		%	Cumulative %
Comp1	3.38	28.10	28.10
Comp2	1.60	13.30	41.40
Comp3	1.17	9.80	51.20
Comp4	0.97	8.10	59.30
Comp5	0.88	7.30	66.60
Comp6	0.84	7.00	73.70
Comp7	0.78	6.50	80.20
...	...	...	...

Table B.1: Principal component analysis Eigenvalue and variance explained

The Eigenvalues are reported in Table B.1 and Figure B.1. The first three components have Eigenvalues greater than 1 (3.38, 1.6 and 1.17). It means that the data suggests a lack of unidimensionality. According to (Gorsuch 1983), the rule of Eigenvalues larger than 1 is applicable when there are fewer than 40 items, the sample size

is large, and the number of factors is expected to be between the number of items divided by 5 and 3. We satisfied all three requirements, because we have 12 items, a large sample size, and the expected number of factors is 4, which falls between 12/5 (2.4) and 12/3 (4). As a result, we suggest that our survey may have two or three components based on Eigenvalues >1.

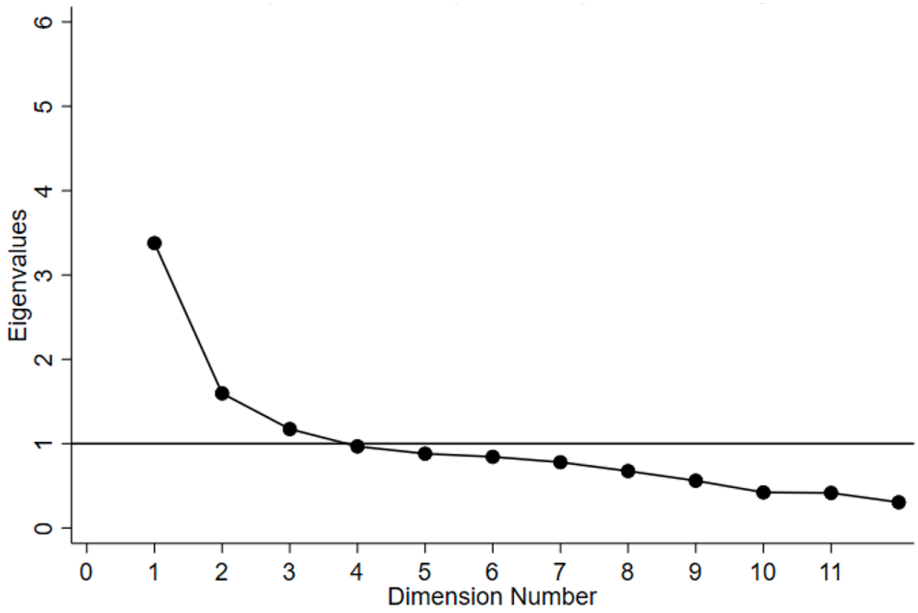


Figure B.1: Scree plot for Principal Component Analysis

B.0.2 Exploratory Factor Analysis

Exploratory Factor Analysis (EFA) is also used to investigate the dimensionality and the possible number of dimensions (assuming the underlying latent traits are normally distributed). The results of the EFA somewhat supported the hypothesized factors stated by Scioli et al., (2011). For the Hope Scale, a five-factor solution emerged, revealing four factors of hope and extracting all negative items under one factor. The factor loadings from Table 7.2 represent the degree of variability in the independent variable that the factor explains. Items 5, 9 and 12 have the lowest loadings and seem not to be strongly related to the factors.

	Factor				
Items	1	2	3	4	5
Attachment					
1. I feel loved by someone (State)		0.63			
2. There are people in my life that I completely trust (Trait)		0.62			
Mastery					
3.I'm making progress towards important goals (State)			0.56		
4. I have a purpose in life (Trait)			0.59		
Survival					
5.I can handle any current or future difficulties. (State)					0.39
6. The future will bring opportunities for a better life. (Mastery/-Survival, Trait)			0.35		0.46
Spirituality					
7. My faith in a higher power gives me the strength to pursue my dreams. (Spirituality, State)	0.69				
8. My spiritual beliefs have empowered me to succeed in life. (Spirituality, Trait)	0.69				
Negative items					
9. I have never felt close to any kind of spiritual force or presence. (Spirituality, Trait)				0.36	
10. I have doubts about achieving those things that really matter to me. (Mastery, Trait)				0.44	
11. I'm running out of options for improving my life. (Survival, State)				0.51	
12. I worry that someone may betray me. (Survival, State)				0.32	

Table B.2: Exploratory Factor Analyses

The two-factor solution separated the eight positive items into a first factor and the four negatively affected items into a second factor.

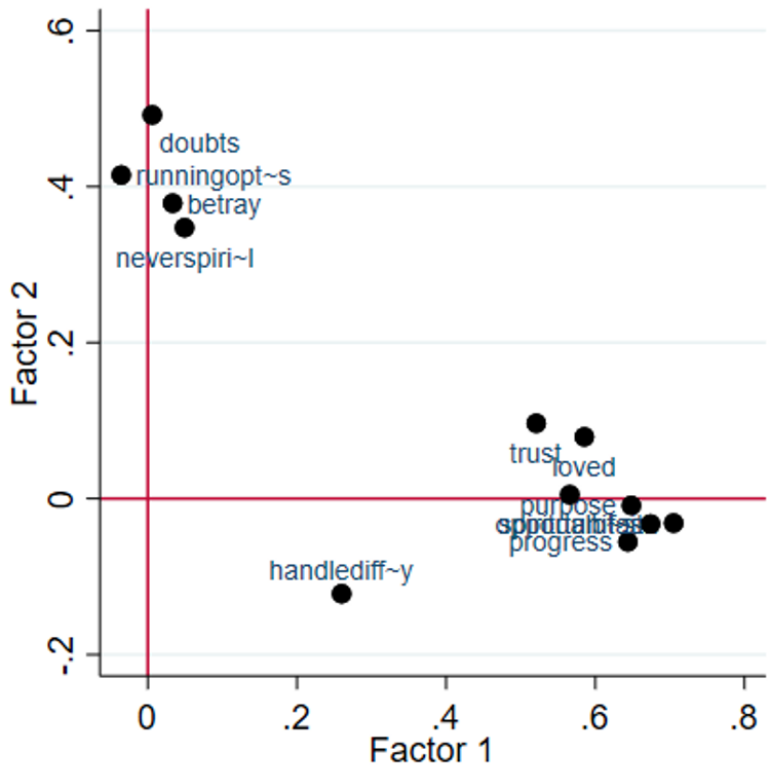


Figure B.2: Factor loadings

The next step is to run omega for all 12 hope items. Omega estimates reliability and also conducts exploratory factor analysis. The result of omega shows the hope survey has 3 factors.

	General factor	F1*	F2*	F3*
1. I feel loved by someone	0.81			
2. There are people in my life that I completely trust	0.67			
3. I'm making progress towards important goals	0.46	0.45		
4. I have a purpose in life	0.44	0.47		
5. I can handle any current or future difficulties		0.3		
6. The future will bring opportunities for a better life	0.3	0.55		
7. My faith in a higher power gives me the strength to pursue my dreams	0.42	0.59		
8. My spiritual beliefs have empowered me to succeed in life	0.4	0.56		
9. I have never felt close to any kind of spiritual force or presence				0.31
10. I have doubts about achieving those things that really matter to me				-0.65
11. I'm running out of options for improving my life				-0.46
12. I worry that someone may betray me				0.35
With eigenvalues of:				
	General factor	F1*	F2*	F3*
	1.98	1.49	0.01	0.86

Table B.3: Omega for Hope

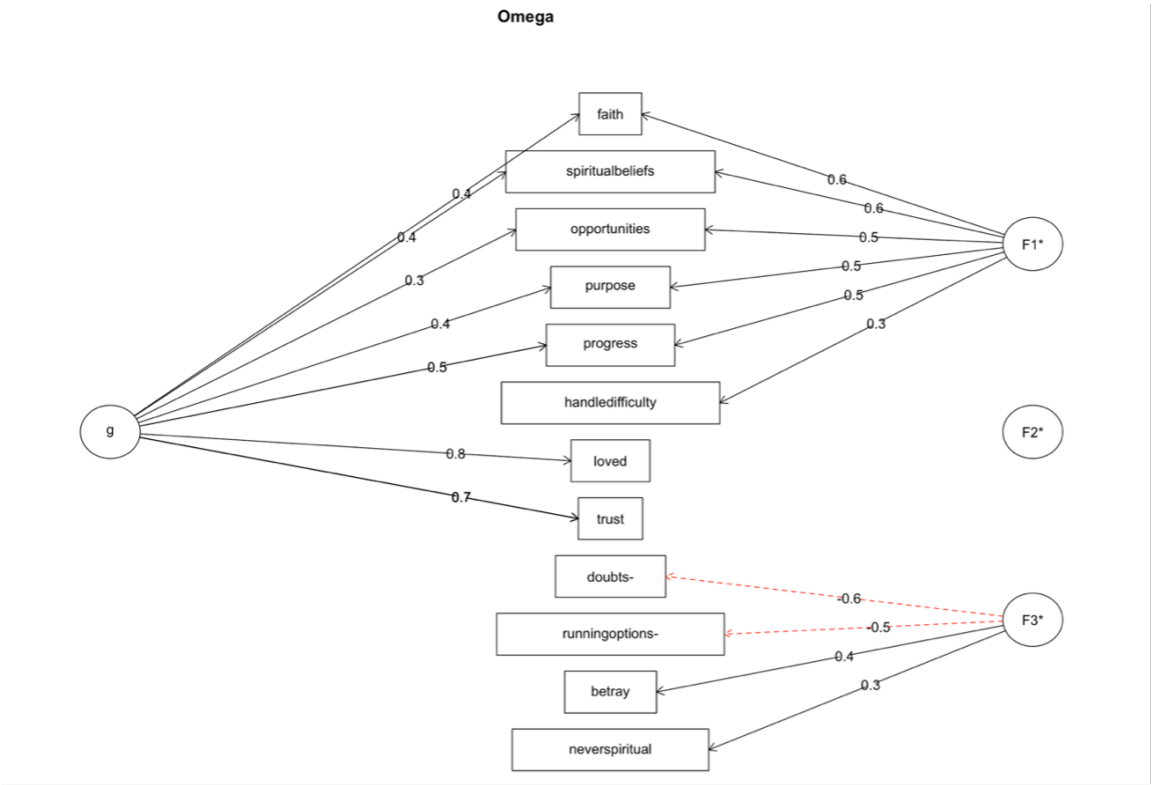


Figure B.3: Omega for all 12 items of Hope

Before testing our models using CFA, let's check whether multidimensionality is critical for the hope we are trying to measure with our survey. Reise et al., (2013) say that when the percent of uncontaminated correlations (PUC) value is >0.80, the overall ECV value has less predictive power; and that when the percentage of uncontaminated correlation (PUC) is <0.80, "researchers may use explained common variance (ECV) values > 0.60 and McDonald's omega hierarchical values >0.70 as provisional benchmarks" to imply that the existence of some multidimensionality is not severe enough to discredit the instrument's primary unidimensional interpretation (Reise et al. 2013). In our case overall McDonald's omega hierarchical is equal to 0.75, and ECV equal to 0.46, PUC equal to 0.67. For positive items PUC=0.75, ECV=0.55, Omega is equal to 0.85. This proves that we might have a multidimensional construct and further analysis comparing unidimensional and multidimensional models is needed.

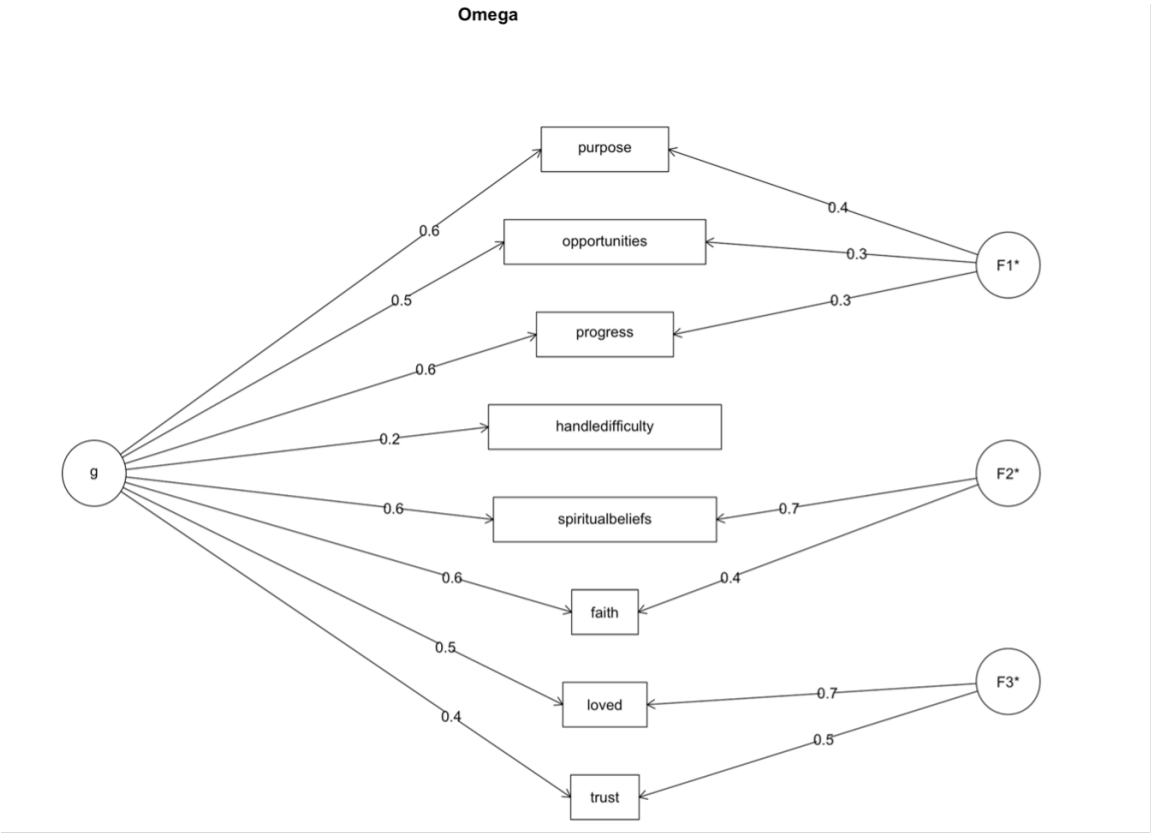


Figure B.4: Omega for 8 positive items of Hope

B.0.3 Confirmatory Factor Analysis

Next, the items were subjected to a single, two and three-factor CFA using the validation sample following the EFA and Omega. We evaluated the confirmatory models' absolute fit using global fit indicators such as the comparative fit index (CFI), the Tucker–Lewis index (TLI), and root mean square error of approximation (RMSEA). We utilised the typical cut - off points of 0.90 or larger for acceptable fit and 0.95 or greater for good fit for the CFI and TLI. Between 0.05 to 0.10, the RMSEA value indicates an adequate fit (Steiger 1990). The results of the CFA can be found in the Table B.4 below.

Confirmatory factor analysis suggests that a 3-factor model with 8 positive items fits the data best (Model 3) based on CFI, TLI and RMSEA. Also, AIC and BIC were computed for all models. It also indicates that Model 3 is a better fitting model, because the smaller the information criteria, the better model–data fit. However, still some research indicates that a little deviation from the assumption of unidimen-

	Unidimensional			**Multidimensional**		
	Model 1	*Model 2*	*Model 3*	*Model 4*	*Model 5*	*Model 6*
	One-factor all positive items	*One-factor all items*	*Two factor all positive*	*Two factor all items*	*Three factor all positive items*	*Three factor all items*
	f1	f1	f1	f1	f1	f1
	opportunities	opportunities	opportunities	opportunities	opportunities	faith
	purpose	purpose	purpose	purpose	purpose	spiritualbelfs
	progress	progress	progress	progress	progress	opportunities
	handlediffculty	handlediffculty	handlediffculty	handlediffculty	handlediffculty	purpose
	faith	faith	faith	faith	f2	progress
	spiritualbelfs	spiritualbelfs	spiritualbelfs	spiritualbelfs	faith	handlediffculty
	loved	loved	f2	loved	spiritualbelfs	f2
	trust	trust	loved	trust	f3	loved
		doubts	trust	f2	loved	trust
		runningoptions		doubts	trust	f3
		betray		runningoptions		doubts
		neverspiritual		betray		runningoptions
				neverspiritual		betray
						neverspiritual
Comparative Fit Index (CFI)	0.77	0.69	0.86	0.77	0.96	0.84
Tucker-Lewis Index (TLI)	0.68	0.62	0.79	0.71	0.94	0.8
Akaike (AIC)	72397	125861	71339	124824	70008	123766
Bayesian (BIC)	72502	126018	71450	124988	70132	123943
RMSEA	0.17	0.12	0.13	0.11	0.07	0.09
SRMR	0.08	0.08	0.06	0.06	0.03	0.05

CFI = comparative fit index (CFI >.90 indicates adequate fit; CFI>.95 indicates very good fit)

TLI = Tucker Lewis index (TLI >.90 indicates adequate fit; TLI >.95 indicates very good fit)

RMSEA = root mean square error of approximation (RMSEA<.10 indicates adequate fit; RMSEA<.05 indicates very good fit)

SRMR: the Standardized Root Mean Square Residual is the square root of the difference between the residuals of the sample covariance matrix and the hypothesised model. It should be <0.08

Table B.4: Confirmatory factor analyses

sionality has little effect on the IRT model and parameter estimations (Dorans and Kingston 1985; Drasgow and Parsons 1983). Thus, as a next step of this research we would recommend checking the loadings for each item in all the models. Even the results support the complex structure of the hope, loadings based on the unidimensional model may slightly differ from the multidimensional models.

Chi-Squared Difference Test							
	Df	AIC	BIC	Chisq	Chisq diff	Df diff	Pr(>Chisq)
Model 5	17	70008	70132	453.96			
Model 3	19	71339	71450	1788.87	1334.91	2	<2e-16 ***
Model 1	20	72398	72503	2849.98	1061.11	1	<2e-16***
Model 6	51	123766	123943	2189.45	-660.53	31	1
Model 4	53	124824	124988	3251.18	1061.74	2	<2e-16***
Model 2	54	125861	126019	4290.76	1039.58	1	<2e-16***
Signif. codes: '***' 0.001 '**' 0.01 '*' 0.05							

Table B.5: Goodness of fit

We compared two unidimensional and four multidimensional models on hope scale for the best fit to our data. The analysis suggests that a three-factor model with eight positive items fits the data best. However, some research indicates that a little deviation from the assumption of unidimensionality has little effect on the IRT model and parameter estimations, thus would advise to conduct a further research on the multidimensionality of the hope scale we used.

Appendix C: Survey questionnaire

O. Hope – Scioli
Please use the following scale to answer each question below.
Questions from this section are asked to all household members aged 10+

	Attachment	**(0) Stronly disagree**	**(1) Disagree**	**(2) Agree**	**(3) Strongly Agree**
KH1601	I feel loved by someone				
KH1602	There are people in my life that I completely trust.				
KH1603	I worry that someone may betray me.				
	Mastery				
KH1604	I'm running out of options for improving my life.				
KH1605	I'm making progress towards important goals				
KH1606	I have a purpose in life.				
	Survival				
KH1607.	I can handle any current or future difficulties.				
KH1608.	The future will bring opportunities for a better life.				
KH1610.	I have doubts about achieving those things that really matter to me.				
	Spirituality				
KH1611.	My faith in a higher power gives me the strength to pursue my dreams.				
KH1612.	My spiritual beliefs have empowered me to succeed in life.				
KH1613.	I have never felt close to any kind of spiritual force or presence.				

Figure C.1: Hope survey questions in English

O. MATUMAINI

Tafadhali tumia kipimo kinachofuata kujibu maswali yafuatayo. Maswali katika sehemu hii yaulizwe kwa wanakaya wenye umri kuanzia miaka 10 na kuendelea.

	Upendo	(0) Nakubali kabisa	(1) Nakubali	(2) Sikubali	(3) Sikubali kabisa
KH1601	Unahisi kupendwa na mtu fulani				
KH1602	Kuna watu katika maisha yako ambao unawaamini kabisa				
KH1603	Una wasiwasi kuna mtu katika familia au jamii anaweza kukusaliti				
	Nguvu ya kujitawala				
KH1604	Unakosa uwezekano wa kuboresha maisha yako?				
KH1605	Unaendeleza na kujitahidi kuyafikia malengo yako ya msingi				
KH1606	Una kusudi mathubuti kwenye maisha yako				
	Kuhimili maisha				
KH1607.	Unaweza kuhimili/kukabiliana na ugumu/changamoto zilizopo na zijazo pia				
KH1608.	Muda ujao utaleta fursa kwa maisha mazuri/bora				
KH1610.	Una mashaka kuhusu kufanikisha mambo yanayokuhusu				
	Imani/hali ya kidini				
KH1611.	Imani yaKo hukujaza nguvu ya kutimiza ndoto zako				
KH1612.	Imani yako ya kidini inakuwezesha kufanikiwa katika maisha yako				
KH1613.	Huamini uwepo wa nguvu za kidini na imani zingine				

Figure C.2: Hope survey questions in Kiswahili

Ask all respondents aged 10+ (Section G-P)

G: Self-efficacy (General Self Efficacy Scale)
Now, I would like to hear your belief in your own abilities as it pertains to dealing with various situations. I am going to read out some statements. For each statement, tell me if it is exactly true (4), moderately true (3), hardly true (2) or not at all (1)

SE1201	I have the skills and knowledge I need to solve difficult problems.	
SE1202	I know how to handle unforeseen problems.	
SE1203	I am able to succeed in ways that really matter to me.	
SE1204	I am capable of finding support from others when I need it.	
SE1205	I draw inspiration from my spiritual beliefs.	
SE1206	I am confident that I can participate in community activities.	
SE1207	I am confident that I can contribute to solutions faced by my community.	
SE1208	There are people in my life that I can completely trust.	
SE1209	The things I need to solve my problems are readily available to me.	

Figure C.3: Self-Efficacy survey questions in English

Waulize wahojiwa wote wenye umri wa miaka 10 na kuendelea (sehemu ya G-P)

G: UFANISI BINAFSI (KIWANGO CHA UFANISI WA KUJITEGEMEA)

Sasa, ningependa kusikia imani yako katika uwezo wako wa kushughulika na hali mbalimbali. Nitakusomea baadhi ya taarifa. Kwa kila taarifa, Niambie kama ni kweli kabisa (4), ukweli kiasi (3), si kweli (2) au si kweli kabisa (1)

SE1201	Una ujuzi na maarifa unaohitajika kutatua matatizo magumu.	
SE1202	Unajua jinsi ya kushughulikia matatizo yasiyotarajiwa.	
SE1203	Una uwezo wa kufanikiwa katika njia zako mwenyewe.	
SE1204	Una uwezo wa kutafuta msaada kutoka kwa wengine pale Unapouhitaji.	
SE1205	Unapata hamasa/msukumo kutokana na imani yangu ya kidini	
SE1206	Una ujasiri ambao unaweza kushiriki kikamilifu katika shughuli za kijamii	
SE1207	Una uwezo wa kuchangia kuhusu suluhu ya matatizo ya jamii.	
SE1208	Kuna watu katika maisha yako ambao ninawaamini kabisa.	
SE1209	Vitu ambavyo ninahitaji kutatua matatizo yangu yote yapo.	

Figure C.4: Self-Efficacy survey questions in Kiswahili

H: Social cohesion, community ownership and trust: Now, I would like to hear your opinions about your community. I am going to read out some statements. For each statement, tell me if you strongly agree (1), agree (2), disagree (3) or strongly disagree (4) with the following statements.		
SE1210	People in this community readily help each other.	
SE1211	People in this community tend to trust each another.	
SE1212	People in this community actively care for people outside of their family who are poor, weak or vulnerable.	
SE1213	People in this community work together to improve the well-being of all children.	
SE1214	The activities that are being implemented in this community to improve the well-being of children are the most important.	
SE1215	People in this community wait for government or NGO support in times of need.	
SE1216	Women take leadership roles in our community councils, or in community groups.	
SE1217	Our leaders listen to input from everyone in the community, including the most vulnerable groups, when making a decision.	
SE1218	Our leaders work hard to improving the well-being of children in this community	
SE1219	Our leaders are able to obtain assistance from outside the community to improve the lives of our children.	
SE1220	In this community, when children are malnourished, we wait for the government to solve the problem.	
SE1221	In this community, if a school building needs some repairs, we wait for other outsiders to come and fix it.	
SE1222	In this community, when children drop out of school, we are not able to do anything about it.	
SE1223	When conflicts arise among the members of this community, we are usually able to solve the problem.	
SE1224	People in my community have close friends of different religions.	

Figure C.5: Social Cohesion survey questions in English

H: MSHIKAMANO WA JAMII NA KUAMINIANA. Sasa, ningependa kusikia maoni yako kuhusu jamii yako. Nitakusomea baadhi ya taarifa. Kwa kila taarifa, Niambie kama nakubali kabisa (1), nakubali (2), sikubali (3) au sikubali kabisa (4) kwa taarifa zifuatazo.		
SE1210	Watu katika jamii hii wanasaidiana.	
SE1211	Watu katika jamii hii wanaonekana kuaminiana.	
SE1212	Watu katika jamii hii wanawajali watu wa nje ya familia zao ambao ni maskini, wadhaifu au wenye mazingira magumu.	
SE1213	Watu katika jamii hii wanafanya kazi kwa pamoja ili kuboresha ustawi wa watoto wote.	
SE1214	Shughuli ambazo zinatekelezwa katika jamii hii ili kuboresha ustawi wa watoto ni muhimu sana.	
SE1215	Watu katika jamii hii wanasubiri msaada wa serikali au shirika lisilo la kiserikali (NGO) wakati wanapohitaji msaada.	
SE1216	Wanawake huchukua nafasi za uongozi kwenye mabaraza au vikundi mbalimbali kwenye jamii	
SE1217	Viongozi wetu kwenye jamii husikiliza mawazo ya mwanajamii yeyote ikiwa ni pamoja na watu walio kwenye mazingira hatarishi kabla ya kufanya maamuzi.	
SE1218	Viongozi wa jamii yetu wanajitahidi kufanya kazi kwa bidii ili kuboresha ustawi wa watoto kwenye jamii	
SE1219	Viongozi wetu wana utayari na uhiyari wa kuhitaji msaada kutoka jamii zingine ili kuboresha ustawi wa watoto wetu	
SE1220	Kwenye hii jamii yetu, watoto wanapokosa lishe huwa tunangojea serikali itatue hilo tatizo	
SE1221	Kwenye hii jamii, pindi majengo ya shule yanapohitaji ukarabati huwa tunasubili msaada kutoka kwa watu wengine wafanye marekebisho	
SE1222	Kwenye jamii yetu hatuna uwezo wa kufanya chochote kwa ajili ya watoto wanaoacha shule	
SE1223	Pindi panapotokea mfarakano/ugomvi baina ya wanajamii wenzetu huwa tuna uwezo na ushirikiano katika kutatua huo ugomvi	
SE1224	Kwenye jamii yetu, Kuna mahusiano mazuri kati ya watu wenye dini tofauti	

Figure C.6: Social Cohesion survey questions in Kiswahili