Journal of International Cooperation in Education

We're all just looking at the stars; how behavioral economics helps us understand the barriers to education programming in Tanzania

Received 22 December 2023
Revised 21 April 2024
24 May 2024
Accepted 12 June 2024

Simon King
Institute of Education, University College London, London, UK, and
Amber Gove
RTI International, Washington, District of Columbia, USA

Abstract

Purpose – We're all just looking at the stars; how behavioral economics helps us understand the barriers to education programming in Tanzania.

Design/methodology/approach – This article uses a qualitative approach to explore the behaviorally normed barriers to quality classroom instruction that contribute towards low learning outcomes. Themed text analysis was applied to qualitative secondary data from seventeen classroom observations and teacher interviews collected from low-performing schools in rural Tanzania.

Findings – It was found that teachers in poor-performing schools in Tanzania were focused on the delivery of curriculum and pedagogy, with a misplaced belief that their pupils were performing adequately. The study found no evidence of teacher resistance to change; instead, the teachers were content and often happy to implement the reading program, believing that teaching phonics-based instruction improved their teaching approach. Teachers sought confirmation of their quality instructional practice from convenient yet inaccurate sources that did not include effective pupil assessment.

Research limitations/implications – As a result of the chosen research approach, findings may lack generalizability.

Practical implications – While existing models of teacher change rely on logic and reason for decision-making, this paper provides evidence that teacher models of change are much more complex and irrational, aligned more closely with insights from behavioral economics (BE). Additionally, this paper justifies that traditional research frameworks that study what works provide an incomplete picture to support effective program improvement.

Originality/value – The application of behavioral economics to research and education programming focused on reducing the restraining forces rather than pushing incentives and other program components.

Keywords Tanzania, Behavioral economics, Diffusion, Education systems, Early grade reading

Paper type Research paper

Background

What works in education research

Bright spots. Positive deviants. Early adopters. Given the same resources, these are different words used to describe individuals whose uncommon behavior helps them solve problems in

Journal of International Cooperation in Education
Emerald Publishing Limited
e-ISSN: 2755-0303
p-ISSN: 2755-029X
DOI 10.1108/JICE-12-2023-0043

the same environment as their peers. In education research, positive deviance is often used to define individuals who can apply an educational innovation while those around them do not, given the same circumstances. Focusing on areas of success is a very common approach when evaluating education programs. For example, UNICEF Innocenti (InnocentiUNICEF Office of Research, 2019) uses positive deviance research to understand *what works* due to positive behaviors and how this can be replicated elsewhere. The Gates Foundation focuses on *what works* (Gates Foundation, n.d.) to improve learning at scale and advise others on designing and implementing successful education programs. In 2023, the United Kingdom's Foreign Commonwealth and Development Office (FCDO) and other donors invested in a *What Works* Hub for Global Education (Blavatnik School of Government, 2023) focused on the science of implementation.

Focusing on *what works* and the positive aspects of where programs are effective is quite appealing; it feels quite intuitive and usually results in optimistic messaging, presenting a hopeful picture of educational progress in lower- and middle-income countries (LMICs).

Researchers frequently use generalized measures to evaluate the impact of education programs, such as standardized means (effect size). However, by their nature, these tools represent a loss of information. If a program has a generalized impact on learning, does it mean all schools under the program are improving? An additional study of data from Early Grade Reading programs by individual school sites suggests that 80% of the programming impact is only explained by approximately 15%-36% of the schools (King & Gove, 2023). The rest of the program schools, which are the majority, have little or no impact on learning.

Why traditional education research paints an incomplete picture

It seems quite logical that education research should focus on discovering what works in the small percentage of schools successfully implementing and scaling these approaches and solutions in all schools. Given that successful schools can implement in a similar environment to schools that do not improve learning, finding out *what works* and introducing these approaches in all schools seems very intuitive. However, there is a fundamental problem with this approach that is hiding in plain sight.

In research, we tend to focus on observed *actions* as an example of how education interventions can be effectively implemented. However, what is usually missed is not what stakeholders are doing but rather *why* they are doing what they are doing. What motivates individuals in successful schools? To understand this, we must consider why some individuals implement or adopt an innovation correctly. (Rogers, 2003, p. 13) describes innovation as "*An innovation is an idea, practice, or object that is perceived as new by an individual,*" a characteristic shared with new education programming as it is introduced.

So, why do the individuals who adopt innovation not follow the majority's normed behavior? The common thread is personality characteristics. These individuals who initially correctly implement or adopt programming, according to Rogers (2003), generally have more ability to deal with abstraction, are more rational, have a higher degree of social mobility, are more intelligent, and can deal more efficiently with risk or uncertainty. This phenomenon was also observed in schools' early-grade reading programming in Nepal, which significantly improved pupil learning (King & Gove, 2023).

So, generally *what works* research shows that an intervention can work in the right conditions and the right individuals. However, scaling positive personality characteristics is not practical or feasible. So, other what options do we have to implement education programs at scale?

The behavioral economics perspective

Behavioral Economist and Nobel Laureate awardee Kahneman (2011) suggests that to facilitate sound decision-making, we should focus on diminishing the restraining forces or

barriers to implementation – not increasing the driving forces. Kahneman credits much of his theoretical thinking to Lewin (1997), who suggested that to achieve behavior change, there is both a wrong and a right way to do it. The wrong way is to increase the driving forces, such as incentives, while the right way is to diminish the restraining forces. This approach, according to Kahneman (2011), is profoundly counter-intuitive. Kahneman and Lewin both suggest that unless the environment is conducive to change, the impact will always be modest to none. This suggests a critical gap in education research and how most education research is designed. By focusing on *what works*, we have little to no idea why something did *not* work in most schools.

Suppose a small percentage of individuals with positive personality characteristics explain the impact of an intervention. In that case, we need to look at typical human behavior for the majority of individuals where there was no impact and figure out why the environment and human behavior prevented positive change. According to Fullan (2015, p. 9), "*. . . the holy grail of change is to know under what conditions hordes of people become motivated to change. The answer is not as straightforward as we would like"*. In other words, by looking at *what works*, we focus on the minority rather than the majority for which we want to see change.

In summary, adding a behavioral economics research perspective achieves two important aims:

(1) It studies barriers to positive change rather than *what works*, and

(2) It applies a behavioral lens to help researchers and others better understand stakeholder actions.

This study conducted a secondary analysis of qualitative data collected in schools in rural Tanzania, focusing on identifying barriers to improved pupil learning. By applying the behavioral science research framework to classroom observations and teacher interviews, we studied behavioral barriers to effective classroom instruction. This approach aligns with other research suggesting that education programs can improve impact by focusing on equity issues and where the effect is poor (Rolleston *et al.*, 2021).

Research question

Research Question: What influences teachers' response to the introduction of an Early Grade Reading Program?

The research question will be answered using a combination of:

(1) Diffusion of Innovations Theory, and

(2) Behavioral Economics

Literature review

Background on early grade reading in Tanzania

In 2016, the Tanzania education system switched to a phonics-based approach from the whole language approach, aligned with the 2016 introduction of the USAID Tusome Pamoja (Let Us Read Together) early-grade reading activity. This shift was prompted in part by evidence from the USAID-funded National Baseline Assessment for the 3Rs (USAID, 2013), the introduction of the Tusome program, and the subsequent National Assessment (an EGRA) conducted in 2016.

Tusome Pamoja was implemented in Zanzibar, Mtwara, Morogoro, Ruvuma, and Iringa (USAID, 2018c). It was designed to support local entities such as the President's Office-Regional Administration and Local Government (PO-RALG) responsible for education

administration build upon existing country initiatives in reading instruction and leverage and improve the existing education system.

Tusome Pamoja had three objectives to achieve the overall aim of improved early-grade reading outcomes: improved classroom instruction, improved performance of management throughout the education system, including teacher monitoring, school leadership, and data for decision-making, and improved engagement of parents and communities. USAID's Jifunze Uelewe (JU) ("Learn and Understand") was the follow-on program that began in 2021. JU had three objectives linked to improved learning outcomes: improved classroom instruction strengthened local and regional ability to support improved pupil learning, and improved community support for safe and inclusive education (USAID, 2021).

JU leverages the Tanzanian education system for program implementation, including using the Ward Education Officer (WEO) as a one-on-one teacher coach, observing individual classes, and providing teachers with feedback on their instructional approach.

Early grade reading programs

EGR Programs are interventions designed to improve foundational literacy skills for primary school pupils in the early grades (typically years 1–3). While early reading interventions long predate the Early Grade Reading Assessment (EGRA) (Dubeck & Gove, 2015), their growth as part of international development and assistance programming grew following the USAID's 2011 Education Strategy and adoption of Sustainable Development Goals (SDG) Indicator 4.1.1, which calls on countries to report on the "*proportion of children and young people: (1) in grades 2/3; (2) at the end of primary; and (3) at the end of lower secondary achieving at least a minimum proficiency level in (1) reading and (2) mathematics, by sex*" (United Nations, 2019).

Spurred by the rapid expansion of early-grade reading assessments (USAID, 2016), EGR programs supported, with varying degrees of emphasis and success, the development of structured teachers' guides, pupil textbooks, and workbooks (Bulat, Dubeck, Green, Harden, Henny, & Sitabkhan, 2017), in-service teacher professional development and coaching, regular formative and summative assessments, increased attention to time on task, and examination of mother tongue/local language as a critical factor in acquiring foundational literacy skills.

In terms of effect size, many EGR programs have been successful (USAID, 2018b). However, in terms of the increase in the percentage of pupils who are proficient readers aligned with SDG indicator 4.1.1(a), the results have been disappointing (Piper & Stern, 2019). A meta-study of EGR program data showed that the average gain associated with early-grade reading programs increased to just three correct words per minute (Sandefur, Alvares de Azevedo, Ju, & Thi, 2023).

The rapid rise of the EGR Programming attracted critics including academic concerns about the emphasis on phonics and oral reading fluency as the core of the testing approach (Hoffman, 2012; Bartlett, Dowd, & Jonason, 2015). Hoffman (2012) was an early critic of structured Early Grade Reading programs who commented that phonics failed to promote reading achievement in the United States and predicted that EGR Programs will have initial, short-term success, which will mask "*. . . the inevitable failure of a commitment to a single method and the dangers inherent in a narrow, technical version*" Hoffman (2012, p. 344). In other words, a single, packaged solution to early-grade literacy fails to account for the need for local conditions, understanding, and decision-making of literacy practices.

Theoretical framework

This study explores teachers' responses to the introduction of EGR Programming by aligning two theories for the research framework: Behavioral Economics (BE) and Diffusion of

Innovations Theory (DOI). For example, DOI discusses how the adoption of an innovation is fundamentally emotional (Rogers, 2003), but it does not fully expand on this idea. BE helps fill this gap, providing a model of how we can study and measure irrational human responses to change (Gordon, 2011). This combined framework also places the focus on improved implementation, rather than concluding the issue is teacher deficiency. For example, Rogers (2003) expects early implementors to be more rational than later adopters but makes it clear that the groups' response should be expected without blame on the individual. This dovetails excellently with BE, where Dan Ariely (2009) says that irrationality of decision-making should be expected, predictable, and can be accounted for in the design.

Diffusion of innovations theory (DOI)

DOI has been applied in over 5,000 studies ("What's so Wrong with the Diffusion of Innovations Theory?" 2021). It is a complex, layered theory that has many different components. For relevance to this research, the two main DOI components used will be: perceived characteristics of innovations, and the innovation-decision process (Rogers, 2003).

Rogers provides five categories by which an individual perceives an innovation during their interaction and ultimately decides to adopt (or not): relative advantage, compatibility, complexity, trialability, and observability (Rogers, 2003). Others have expanded on this list of characteristics. Moore and Benbasat (1991) identified image (perceived to enhance one's social status), and voluntariness of use as additional innovation characteristics.

DOI theory describes the decision to adopt innovation as a process that takes place over time. While every process is unique to the individual, there are five main stages to the innovation-decision process:

(1) Knowledge - The complexity of the innovation is essential.

(2) Persuasion - focused on individual attitudinal formation and change regarding innovation.

(3) Decision – the individual decides to use the innovation or reject it. This decision is on a scale from rejection, partial adoption/rejection, to adoption.

(4) Implementation – implementation forms include re-adoption, partial adoption, or incorrect performance.

(5) Confirmation – the individual looks for reinforcement of their innovation decision.

Rogers (2003).

Some components of DOI have been applied to education research. For example, the main factors determining the adoption of a specific set of ICT skills for teacher trainers in Cambodian public schools are if the innovation was perceived to be easy to adopt (i.e. complexity) and mandatory (i.e. the voluntariness of use) (Richardson, 2011). Hughes and Keith (1980) also found a statistically significant association between perception of the innovation and teacher implementation.

Behavioral economics

BE pushes against conventionality about how we think about participants in the standard economic model of human behavior, that individuals "... *think and choose unfailingly well*" (Thaler & Sunstein, 2009, p. 6). A presumption of logical reasoning is a common model of human behavior used for EGR Programming which often incorporates Guskey's (2002) model of teacher change. This model suggests teachers logically use pupil learning outcomes to reflect on a new instructional approach and make modifications to the approach before entering a new reflection cycle.

Instead, BE positions that individuals' behavior is irrational yet predictable (Ariely, 2009). While early research in BE was focused on applying psychology and sociology to the economic model to understand better how humans behave, identifying the irrationality of human behavior in economics is not exactly new. John Maynard Keynes acknowledged the existence of "animal spirits" (John Maynard Keynes, 1936) driving financial decisions.

The most common application of BE in education is to understand and/or nudge pupils' behavior in and out of the classroom (Koch, Nafziger, & Nielsen, 2015; Levitt, List, Neckermann, & Sadoff, 2012). However, there has been little traction with BE in education regarding understanding teacher behavior, social norms, and classroom instructional approaches in LMICs (Jabbar, 2011; Levitt *et al.*, 2012; Stevano, 2019).

BE research mainly focuses on civil society. Thaler and Sunstein (2009) proposed that BE be applied using a Libertarian Paternalism (LP) model, where individuals are free to choose. However, how model works within an education system is uncertain, where there are often expectations and accountability of teachers and other stakeholders.

Data source

This study used secondary data analysis of seventeen qualitative classroom observations and teacher interviews conducted in Tanzania. Existing qualitative data collected in March 2022 by RTI International are analyzed using this study's theoretical framework and research question. The secondary data source was a purposeful selection of 12 low-performing schools in the rural Tanzanian Districts of Iringa and Morogoro. These schools were part of the Jifunze Uelewe (JU) EGR Program and the data was collected for operational research. Low-performing schools were originally selected by JU using the results of the Tanzania Primary School Leaving Examination (PSLE) (The National Examinations Council of Tanzania, 2021). Seventeen grade 1 and grade 2 teachers from twelve schools were observed teaching and interviewed immediately after. Other teachers were not available or present at the school during the data collection visit. The classroom observation was recorded as a text narrative and the interview data consisted of primary questions with follow-up secondary questions. The classroom observation notes were translated into English and entered into a template. The teacher interview audio recordings were translated and transcribed into English. The teacher interview and classroom observation data were linked by teacher. The secondary data was deemed appropriate for the research framework, is it helped build a reasonable picture of the environmental influences behind teacher decision-making (Gordon, 2011).

For example, the teachers were asked. "*Has clear information been expressed to you regarding the level of achievement or progress pupils in your class should show by the end of the year?*". Secondary questions asked:

(1) *"What are those expectations and whose responsibility is it to achieve these expectations?"*

(2) *"Do you believe these expectations are achievable? How do you know?"*

These prompts were then cross-referenced with the teacher's use of assessment and interaction with students during the classroom observation.

Methodology/data analysis

The data were analyzed using Nvivo 12. The data was analyzed using the six-stage approach to thematic analysis (Braun & Clarke, 2006), using a mostly deductive deductive approach to align and code themes from the theoretical framework. These themes included, but were not

limited to characteristics of adopters (Rogers, 2003), a model of teacher change (Guskey, 2002), and power of the default (Thaler & Sunstein, 2009).

Consideration was also given to an inductive theme approach; building new themes that were not included as part of the deductive themes (Braun, 2022). An example of a deductive theme developed was bounded rationality (Simon, 1957).

Limitat`ions

The 5th-grade PSLE achievement reading data was used to select low-performing schools for the secondary data collection. However, the participants in the study are teachers of grades 1 and 2. As PSLE is conducted with grade 5 pupils, it should not be inferred that achievement levels will be similar in grades 1 and 2. Consequently, the study does not fully account for variation in learning outcomes by grade within schools.

Ethics

All publicly identifiable information (PII) from the secondary data was not used in the dissemination of the findings of this study. All secondary data was kept securely behind a firewall.

Results

RQ. What influences teachers' response to the introduction of an Early Grade Reading Program?

A key aspect of the secondary data analyzed is that teacher response was assessed after six years of early-grade reading programming. The data did not indicate how long individual teachers had been involved in the early-grade reading programs. However, it was determined that all had undergone an interaction with the intervention as they were all observed in class using the program materials. These two key findings help frame the other findings.

Theme – cognitive dissonance: teachers believed that most or all their pupils would be proficient readers by the end of the year

This study was designed with a qualitative approach. Secondary data from low-performing schools were analyzed. The teachers responded that most or all their pupils would be proficient readers by the end of the year. When asked if their pupils would achieve this level of proficiency by the end of the year, they replied that most or all would be able to. In comparison, The Tanzania National Early Grade Reading Assessment (USAID, 2018a) reported the percentage of pupils in grade 2 who can read thirty or more correct words per minute as 36.1%. The suggestion that the teachers in this study did not know the reading skill level of performance of their pupils is not a unique finding. These findings align with research conducted by Twaweza (2023) who mentions that teachers in Tanzania are either unaware or do not acknowledge student literacy issues. These findings also are consistent with a study of teachers in Southeast Asia, suggesting that teachers were mostly unaware of the reading level of their pupils, particularly those who were low-achieving (Djaker, Ganimian, & Sabarwal, 2022).

Theme – perception of the innovation: resistance to reading the reading program was neither observed in the classroom nor expressed during the teacher interview

Aligned with other educational research regarding why teachers adopt new programming (Richardson, 2011), teachers suggested implementing the new program was mandatory. However, the teachers' interpretation of mandatory focused not only on accountability but

also on the idea that it was mandatory because, as one teacher said, it made *". . . learning for these young children more meaningful*". In other words, many teachers believed using the program materials was mandatory because they felt obliged to teach a program superior to their previous teaching approach before the program introduction. All the teachers thought the reading program was superior to their prior approach. One teacher commented:

> I feel good because this new approach makes my lesson easy for pupils to understand. For example, using the I do, we do, and you do model gives pupils a chance to listen, practice together, and later practice themselves. So, these approaches are better than what I had before.

Most teachers said they did not find the new reading program too difficult to adopt, and there were clear indications of motivation to learn to implement the new program. One teacher said, "*You know if something is benefiting you, you will try your best to ensure you use it.*". Another said, "*For me, it was easy because I had the passion to learn new approaches*."

The teachers were all observed implementing the phonics-based literacy content, demonstrating a competent understanding of the content and pedagogical approach. A few teachers struggled a little with teaching sounding the correct phonemes (letter sounds), but not to the point of poor delivery of content and something that could be corrected with some support.

Consequently, the teachers were interviewed late in Rogers' (2003) innovation-adoption process; based on classroom observation and the positive response of the teachers to the program, they were likely at the confirmation stage, given their feedback regarding the program.

If teacher resistance to the new program is not an issue, why are the pupil learning outcomes in these schools so poor?

All teachers were observed modifying the reading program approach in a specific way. The program lesson plans followed an "*I do, we do, you do"* gradual release approach (Gove & Cveslich, 2011). This approach is where teachers introduce the new skill or content, pupils practice or repeat the latest content, and finally, the pupils practice the new skill or content in groups and individually. Most of the teachers adequately delivered the first two aspects of the lesson plan. They were all familiar with introducing the phonics-based content as designed and had pupils repeating or individually answering questions during "*we do*" the whole class interaction. However, their application of the third part of the class, the pupils practicing and mastering the new skills, was poorly executed.

The observed modification to the reading program occurred when the pupils were required to practice the new content or skill. All the teachers modified the approach in a very similar way. Some teachers had the pupils work in groups; others were instructed to work individually or in pairs to practice the new skills and write them into their exercise books. Teachers would move around the classroom, checking and marking exercise books. However, teachers mostly checked that pupils had copied what was on the blackboard into their exercise books. It was observed a few times that if a pupil could not complete the work, they did not ask for support from the teacher; instead, they waited to see if the teacher would come to them. Many pupils were observed without exercise books and utensils to write with. Either they borrowed from their peers or did not get the opportunity to write. Finally, this aspect of the class was often shortened by the lesson plan timing it should have had. So, pupil quality time on task was lacking in every classroom observation. The framework has possible reasons to explain why this was the case. Firstly, it could be a simple case of mental fatigue. The teacher has been the focal point of the class up to the final part of the class, the pupil practice. Teacher mental fatigue will only increase during a 45-min class, so does Type I thinking (fast, emotional, and instinctive) take over Type II (slow, logical, deliberate)

thinking? Another explanation could be the power of the default; teachers often struggle to be effective when they focus on delivery over pupil learning.

Theme – mental effort: teachers are comfortable and often familiar with adopting new curricula

It seems the teachers in the study can adopt the program and change it in limited ways. Teachers seem to focus on the delivery aspects of class; they seem able to learn new curricula and introduce new content using new pedagogical practices, but this is focused delivery. The missing component that appears to be an understanding of improvement is pupil learning.

Models of teacher change usually mention pupil assessment as how teachers assess if their change in instructional practice is effective (Guskey, 2002). Rogers (2003) called this stage *confirmation*, where the individual has adopted the innovation and now looks for evidence to confirm that their adoption of the innovation was a good decision. Rogers claims this decision is fundamentally emotional, whereas Guskey's (2002) model relies on rationality. Teachers draw on different stimuli to confirm that their new teaching approach is effective, but not pupil quality assessment. One teacher said, "*I teach my pupils in a way that they all get involved and enjoy my lesson, and by the end of the year, they are all able to read.*" This type of response, highlighting interaction with their pupils, was a common response from teachers, suggesting that teacher confirmation of an effective instructional approach is gained from positive interaction with the pupils. Additionally, the analysis suggestion is also quite possible that many teachers get confirmation by the fact that they are delivering the new curriculum. As one teacher said, "[the phonics approach] . . . *helps pupils to learning in more effective way compared to old approaches*". However, we know that effective instructional practice has many components other than just curriculum.

Theme - education system - support and set expectations

The teachers interviewed could describe many systems and mechanisms to support learning. The Ward Education Officer (WEO) and head teacher usually provide a support role within the education system in Tanzania. However, teachers reported getting support from many sources, including the head teacher, peer teachers, and lead teachers. However, importantly, every teacher described getting support. All the teachers expressed a reflective process for improvement; they were observed and given goals/pointers to improve on.

Many systems support guidelines suggest setting expectations for systems (RTI International, 2018). In Tanzania, it was agreed and communicated (through the systems cascade) that by the end of grade 2, pupils should meet the grade level benchmark of thirty correct words per minute. All the grade 2 teachers interviewed correctly recalled this expectation.

Setting expectations for learning outcomes in an education system seems logical and reasonable. It is not known if the teachers had a degree of cognitive dissonance; they have seen assessments of their pupil's reading skills and deny the results.

Discussion

Teacher resistance was not a problem

This study showed that teachers were content to implement the reading program after six years of early-grade reading programming. More than that, for many teachers, "mandatory" was their moral obligation to teach using what they considered a superior reading program. Given the all-too-common alternative's lack of teaching and learning resources, why would they not want to implement the new program? Indeed, most of the teachers positively perceived the program and believed they were implementing it effectively.

The challenge of doing anything other than teaching the program is considerable. Given that the program resources are often the only teaching and learning materials a teacher has, not teaching would mean a lack of compliance.

This lack of resistance uncovered by this study is not necessarily new in other countries and contexts. According to the Brookings Institute, no substantial evidence exists that Common Core positively impacted pupil achievement (Loveless, 2021). Brookings detailed key findings to conclude that there was no considerable evidence teachers and other stakeholders resisting the implementation of Common Core (Loveless, 2021). Across borders, while environment and culture make us different, our human nature makes us behave worldwide similarly.

What else might explain the low learning outcomes? Before the introduction of EGR Programming in Tanzania, a key cause of low literacy outcomes was attributed to children not being able to "*. . . decode words effectively and efficiently*" (USAID, 2013, p. 67), however switching to phonics-based instruction still has not significantly addressed learning outcomes. The 2013 National Baseline Assessment also mentions a lack of teaching and learning materials, issues of teacher and student attendance, school readiness, and community involvement as potential issues (USAID, 2013).

Teacher confirmation of instructional practice

Bounded rationality, or specifically *satisficing*, the process of making a decision that is satisfactory rather than optimal (Simon, 1957), is not new in the study of teacher behavior. Generally, it has been studied in Western contexts very different from Sub-Saharan Africa and emphasized positive decision-making but lacks evidence of improved learning outcomes (Ralaingita & du Plessis, 2020). Our study suggests that teachers use stimuli for an emotional confirmation aligned with Rogers (2003) rather than even bounded rationality. It is important to note that a big motivation behind Herbert Simon's (1957) development of bounded rationality is that he was a critic of the economics notion of human behavior, that individuals had unlimited potential to process information to rationalize decision-making (Mullainathan & Thaler, 2000). Further, this study suggests that teachers seek confirmation based on emotional, not limited, rational responses. Poor learning outcomes in schools prove this, but most teachers believe their pupils can read. Additionally, the teacher's perception of their teaching may have contributed. A typical teacher responded, "*I feel lucky because now my teaching has improved, and I also see my pupils improve in learning. They enjoy my approaches, which helps them learn more easily*".

An alternative model of teacher change

The model of teacher change used to design early-grade reading programs tends to follow approaches such as Guskey (2002).model of teacher change (Figure 1).

Guksey describes teacher confirmation that classroom practice is effective through a change in learning outcomes. This model relies on a logical thought process and the availability of pupil learning outcomes. However, while our teachers are near or at the confirmation stage, they are not getting confirmation from reliable assessments of learning outcomes. As discussed, exactly where they are getting confirmation of effective instructional

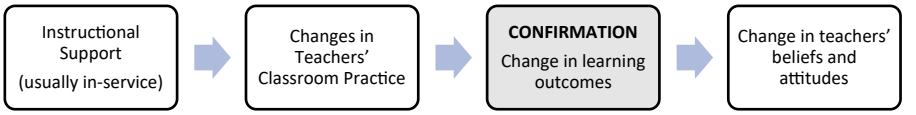


Figure 1.
Guskey's model of teacher change

Source(s): Guskey (2002)

practices is uncertain. Still, given the cognitive dissonance between teacher's belief in learning outcomes and actual learning outcomes, the stimuli for confirmation do not seem to be helpful for instructional practice (Figure 2).

As discussed, Guskey relies on a logical reflective process for confirmation, while DOI theory says this stage is primarily emotional (Rogers, 2003). It seems important to understand this emotional confirmation stage better. For example, if provided with quality inputs that can inform instruction, will teachers recognize that these inputs are necessary and respond rationally? This approach is counter to the foundation of behavioral economics. Kahneman (2011) argues that you cannot make people rational; you can nudge good decision-making through design.

Do education systems focused on compliance encourage poor decision-making?

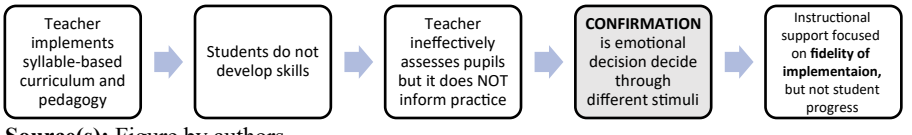
All but one of the seventeen teachers in this study expressed that they were content. The teachers implemented the education program by teaching the scheduled lesson plans on the scheduled days. They believed the program improved their classroom instructional approach. Most reported being observed by head teachers, peers, or WEOs. These are all findings that suggest an education system is functioning adequately. This might be a great confirmatory qualitative success story for *what works* type research. However, this qualitative study was conducted in low-performing schools, thus uncovering contradictions. It seems that a combination of accountability systems and structured programming has resulted in teachers moving through lesson plans with the understanding that this is their key responsibility and that this approach would furnish positive learning outcomes for their pupils.

However, a cycle of events seems to exist that harms the chances of pupils making learning progress. The teacher teaches the lesson plan, but pupils do not develop and understand the key concepts; the teacher seems unaware of this and moves on to the next lesson the following day, as the program guides them to do. So, do teachers receive confirmation that they have achieved professional competence through compliance? Is teaching the lesson plans good enough?

Kahneman (2011) comments that when individuals are asked to improve, they generally *try a little harder*. However, when the barriers to effective instructional practice have not been addressed, this improvement process does not improve outcomes.

Changing curriculum is familiar to teachers, requiring a low mental effort for teachers to make change achievable and reduce stress. Given that all the teachers in the study behaved similarly, this seems to be normed behavior. The schools in the survey were spread out geographically, so it is unclear if the education system has any connection with this normed behavior.

The study's findings resemble an education system akin to Lantt Pritchett's *Isomorphic Mimicry* (Pritchett, 2011), which describes where a country has developed an administrative system of checks and balances. However, he argues this system cannot implement more complex tasks. In other words, stakeholders are doing what they believe is expected of them, but it is not impacting the complexities of improving the quality of instructional practice. Fullan (2015) also mentions the challenge of when school leaders are saddled with the



Source(s): Figure by authors

Figure 2. Proposed model of teacher behavior

responsibility of holding individuals accountable, rather focusing on all the effective drivers of change.

Why do teachers focus on delivery over pupil learning?

The teachers in this study primarily focused on delivering the curriculum through new pedagogical practice at the observed expense of focusing on pupil learning. Exactly why this behavior exists is unclear. There are different possibilities (or combination of) that could be further researched:

(1) Teachers are familiar with curricula and content change. This is teacher delivery and can be quite procedural. They are less familiar with changing their entire instructional approach to being student-centered. This is likely more associated with behavior change.

(2) Structured education programs primarily focus on introducing new curricula and pedagogy and focusing support on teachers improving delivery, giving the impression that if this is prioritized, quality classroom instruction and improved learning outcomes are assured.

(3) Systems often hold teachers accountable for the easy-to-measure fidelity of teaching new curriculum and pedagogy in the classroom, so teachers focus on this aspect and less on whether pupils are progressing.

(4) Type I thinking (Kahneman) influences decision-making:

- Mental fatigue makes teachers tired towards the end of class, which is the time for the key *quality time on task* for pupils, which might explain why it is poorly executed.

(5) Teachers believe that they are exerting all the effort they can (Sabarwal & Abu-Jawdeh, 2018).

Conclusion

The conclusion of this study is twofold. First, that *what works* research provides only a partial understanding if we want to scale education programming successfully. Second, that education programs should not be designed with a presumption of some a high degree of rational participant behavior, and to achieve better implementation, we need to acknowledge that knowledge does not necessarily change practice.

Research the behavioral barriers where impact is low, not just the small pockets of success

This study suggests that education programming design can fail to properly consider how teachers and other stakeholders will likely respond behaviorally. Using our theoretical framework, we see that a focus on delivery at the expense of pupil learning remain a barrier to progress. Moving forward, we recommend researching existing education programs using a behavioral economics approach, looking at where impact is low and uncovering behaviors and contradictions that present barriers to progress.

Design and implement programs based on researched human behavior

In the background section of this study, we acknowledged the ongoing research discourse regarding phonics-based instruction and structured pedagogy. Teachers in this study implemented the program through lesson plans and had a positive perception of the program.

However positive the teacher perception to the new program, it does not necessarily mean that implementation and program design is effective. The research findings from this study are more relevant for re-thinking education systems approaches and program implementation design, as illustrated below.

Diffuse or cascade new approaches?

Most education programs in LMICs are implemented similarly. They leverage the education system cascade of central, district, and school-level programming roll-out. Knowledge is passed down through a series of trainers and teacher trainers training. Then, ongoing support is provided, usually through district-level coaches and head teachers. Unlike small research pilots, most donor-supported programs will leverage these education system structures. However, can we consider diffusion as an alternative? Rogers (2003, p. 5) defines Diffusion of Innovation theory as "*. . . the process in which an innovation is communicated through certain channels over time among the members of a social system*". While we do not suggest that all innovations should be designed and diffused at the school level, a model where education programming is designed at the central level but laterally diffused at the local level seems worth investigating. It seems the weakness in the cascade model is less in the introduction of the literacy program, but more in how it is supported through WEOs as coaches. This problem is confirmed by Thomas Guskey, who states that "*. . . logic, reason, and philosophical arguments*" (Guskey, 2020, p. 19) do not work, but is the approach frequented by education leaders. Additionally, Fullan (2015, p. 43) states, "*The problem is that no nation has improved by focusing on the individual as the driver*".

Diffusion theory (Rogers, 2003) and social norm theory (Bicchieri & Noah, 2017) could be central theories in the design of such an implementation. However, we caution that it is unlikely that forming communities of practice with little focus on behavioral science will be effective. Rogers (2003) suggested that most diffusion practice requires a level of curation. For example, Rogers (2003) writes about the importance of change agents; in our case, peer teachers with effective practice who possess a degree of social influence. This approach can easily be reduced to an ineffective strategy of having teachers watch demonstrations of effective practice without considering social norms. In other words, the implementation design again rests on the presumption of logical human behavior and how knowledge will change practice.

There are already efforts in Tanzania to use influencers or champions to diffuse effective practice and advocacy, such as promotion of girls' education through the Education Champion Network (ECN). Additionally, Uwezo's Uwezo Na Jamii activity (UWEZO, n.d.) uses assessment scores to create awareness of the learning outcomes of children in a community. The model assumes humans respond to logic and reasoning. It would be interesting to find out if community support for improving learning outcomes can be nudged using a model of human irrationality and an approach that combines DOI and social norm theory.

Should we re-think the gradual release model?

Teachers and other systems stakeholders seem to focus on a new curriculum in adopting the reading program. It is the most familiar and least stressful aspect of change that teachers can apply. Teachers are often already quite used to changing content by introducing new textbooks and subjects or teaching a different grade. General habits of instructional practice remain, but more challenging aspects of instructional practice, such as pupil practice of new skills, remain poorly done, as the findings of this study suggest. The Jifunze Ulewe program leverages the gradual release model. There is logic in this as at least the "I do" and "we do" components of the class are already familiar to most teachers, and they perform adequately

on these aspects of classroom instruction. However, given that teachers in this study struggled with lesson timing and likely mental fatigue towards the end of class, the pupil practice at the end of class was generally poorly done. Programs should adjust lesson designs to *make it easy* for teachers to focus on pupil practice by starting classes with this component.

References

Ariely, D. (2009). *Predictably irrational: The hidden forces that shape our decisions*. London: HarperCollins. Rev. and exp.

Bartlett, L., Dowd, A. J., & Jonason, C. (2015). Problematizing early grade reading: Should the post-2015 agenda treasure what is measured?. *International Journal of Educational Development*, *40*(C), 308–14. doi: 10.1016/j.ijedudev.2014.10.002.

Bicchieri, C., & Noah, T. (2017). Applying social norms theory in CATS programming. Available from: https://repository.upenn.edu/pennsong/15

Blavatnik School of Government (2023). New what works Hub for global education will turn evidence on learning into reality for millions of children, July 20, 2023. Available from: https://www.bsg.ox.ac.uk/news/new-what-works-hub-global-education-will-turn-evidence-learning-reality-millions-children

Braun, V. (2022). *Thematic analysis: A practical guide*. London: SAGE Publications.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. doi: 10.1191/1478088706qp063oa.

Bulat, J., Dubeck, M., Green, P., Harden, K., Henny, C., & Sitabkhan, Y. (2017). *What we have learned in the past decade: RTI's approach to early grade literacy instruction*. RTI Press.

Djaker, S., Ganimian, A. J., & Sabarwal, S. (2022). Primary- and middle-school teachers in South Asia overestimate the performance of their students. In *RISE Annual Conference 2022*, England. Oxford. Available from: https://riseprogramme.org/events/rise-annual-conference-2022

Dubeck, M. M., & Gove, A. (2015). The early grade reading assessment (EGRA): Its theoretical foundation, purpose, and limitations. *International Journal of Educational Development*, *40*. doi: 10.1016/j.ijedudev.2014.11.004.

Fullan, M. (2015). *The new meaning of educational change* (5th ed.). New York: Teachers College Press.

Gates Foundation (n.d.). Structured pedagogy series. *Science of Teaching*, (blog). Available from: https://scienceofteaching.site/how-to-guides/learning-outcomes/ (accessed 19 February 2022).

Gordon, W. (2011). Behavioural economics and qualitative research – a marriage made in heaven?. *International Journal of Market Research*, *53*(2), 171–85. doi: 10.2501/IJMR-53-2-171-186.

Gove, A., & Cveslich, P. (2011). *Early reading: Igniting education for all: A report by the early grade learning community of practice | Unesco IIEP learning portal*. Research Triangle Institute. Available from: https://learningportal.iiep.unesco.org/en/library/early-reading-igniting-education-for-all-a-report-by-the-early-grade-learning-community-of

Guskey, R. T. (2002). Professional development and teacher change. *Teachers and Teaching, Theory and Practice*, *8*(3), 381–391.

Guskey, T. R. (2020). Flip the script on change: Experience shapes teachers' attitudes and beliefs. *The Learning Professional*, *41*(2).

Hoffman, J. V. (2012). Standpoints: Why EGRA—a clone of DIBELS—will fail to improve literacy in Africa. *Research in the Teaching of English*, *46*(4), 340–57. doi: 10.58680/rte201219761.

Hughes, A. S., & Keith, J. J. (1980). Teacher perceptions of an innovation and degree of implementation. *Canadian Journal of Education*, *5*(2), 43–51. doi: 10.2307/1494312.

Innocenti, UNICEF Office of Research (2019). *Data must speak*. UNICEF-IRC. Available from: https://www.unicef-irc.org/research/data-must-speak/

Jabbar, H. (2011). The behavioral economics of education: New directions for research - Huriya Jabbar, 2011. *Educational Researcher*, *40*(9), 446–453, December. doi: 10.3102/0013189X11426351.

John Maynard Keynes (1936). *The general theory of employment interest and money*. London: Macmillan.

Kahneman, D. (2011). *Thinking, fast and slow* (1st ed.). New York: Farrar, Straus and Giroux.

King, S., & Gove, A. (2023). Why scaling 'what works' usually doesn't work for foundational literacy programs, (blog). 2023. Available from: https://www.globalpartnership.org/blog/why-scaling-what-works-usually-doesnt-work-foundational-literacy-programs

Koch, A., Nafziger, J., & Nielsen, H. S. (2015). Behavioral economics of education. *Journal of Economic Behavior and Organization*, *115*(July), 3–17. doi: 10.1016/j.jebo.2014.09.005.

Levitt, S. D., List, J. A., Neckermann, S., & Sadoff, S. (2012). The behavioralist goes to school: Leveraging behavioral economics to improve educational performance. National Bureau of Economic Research. Working Paper 18165. Working Paper Series. doi: 10.3386/w18165.

Lewin, K. (1997). *Resolving social conflicts & field theory in social science*. Washington, DC: American Psychological Association.

Loveless, T. (2021). *Why common core failed*. Brookings. (blog). March 18, 2021. Available from: https://www.brookings.edu/blog/brown-center-chalkboard/2021/03/18/why-common-core-failed/

Moore, G. C., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research*, *2*(3), 192–222. doi:10.1287/isre.2.3.192.

Mullainathan, S., & Thaler, R. H. (2000). Behavioral economics. *NBER Working Papers*, NBER Working Papers 7948, October. Available from: https://ideas.repec.org//p/nbr/nberwo/7948.html

Piper, B., & Stern, J. M. B. (2019). Resetting targets: Examining large effect sizes and disappointing benchmark progress, April. doi: 10.3768/rtipress.2019.op.0060.1904.

Pritchett, L. (2011). One size doesn't fit all: Lant Pritchett on Mimicry in development. (blog). March 14, 2011. Available from: https://www.cgdev.org/blog/one-size-doesn%E2%80%99t-fit-all-lant-pritchett-mimicry-development

Ralaingita, W., & du Plessis, J. (2020). "Satisficing" in early grade reading: Applying reasonably good strategies in imperfect contexts. In Wiseman, A. W. (Ed.), *Annual Review of Comparative and International Education 2019* (*39*, pp. 191–208). Leeds: Emerald Publishing.

Richardson, J. W. (2011). Technology adoption in Cambodia: Measuring factors impacting adoption rates. *Journal of International Development*, *23*(5), 697–710. doi: 10.1002/jid.1661.

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). New York: Free Press.

Rolleston, C., Crouch, L., & Gustafsson, M. (2021). Eliminating global learning poverty: The importance of equalities and equity. *International Journal of Educational Development*, *82*. doi: 10.1016/j.ijedudev.2020.102250.

RTI International (2018). *Education policy systems and governance*. RTI Interational. Available from: https://www.rti.org/sites/default/files/related-content-files/idg_education_institutions_0.pdf

Sabarwal, S., & Abu-Jawdeh, M. (2018). What teachers believe : Mental models about accountability, absenteeism, and student learning. *The World Bank*. Available from: http://documents.worldbank.org/curated/en/804301527601436747/What-teachers-believe-mental-models-about-accountability-absenteeism-and-student-learning

Sandefur, J., Alvares de Azevedo, T., Ju, X., & Thi, L. (2023). *Phonics and Foreign aid: Lessons from a decade of USAID early-grade reading evaluations*. Center for Global Development. Available from: https://www.cgdev.org/event/phonics-and-foreign-aid-can-america-teach-world-read

Simon, H. (1957). *Administrative behavior: A study of decision-making processes in administrative organisation* (2nd ed.). New York: Macmillan.

Stevano, S. (2019). The randomistas won the 2019 Nobel prize for economics. Will they also win the debate on development economics?, SOAS Blog (blog). Available from: https://www.soas.ac.uk/blogs/study/randomistas-economics-nobel-prize/ (accessed 16 October 2019).

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth and happiness*. London: Penguin.

The National Examinations Council of Tanzania (2021). *NECTA | PSLE results*. The National Examinations Council of Tanzania. Available from: https://necta.go.tz/psle_results

Twaweza (2023). Delivering quality education? 10 insights from public primary schools in Tanzania. Available from: https://twaweza.org/delivering-quality-education-10-insights-from-public-primary-schools-in-tanzania/ (accessed 20 April 2024).

United Nations (2019). *SDG indicators*. SDG Indicators. December 2019. Available from: https://unstats.un.org/sdgs/metadata/?Text=&Goal=4&Target=4.1

USAID (2013). National baseline assessment for 3Rs (reading, writing, and arithmetic) in Tanzania (summary brief). Available from: https://shared.rti.org/content/national-baseline-assessment-3rs-reading-writing-and-arithmetic-tanzania-summary-brief

USAID (2016). *Early grade reading assessment toolkit* (2nd ed.). United States Agency for International Aid.

USAID (2018a). *ABE-ACR Tanzania national early grade reading assessment (EGRA) final report*. USAID. Available from: https://shared.rti.org/content/abe-acr-tanzania-national-early-grade-reading-assessment-egra-final-report

USAID (2018b). *Early grade reading barometer*. Early Grade Reading Barometer. Available from: https://earlygradereadingbarometer.org/india-start-early-read-in-time-odisha/benchmarks/sankey

USAID (2018c). *Tanzania - Tusome Pamoja*. USAID. Available from: https://www.usaid.gov/sites/default/files/documents/1860/2018-10-29_Tusome_Pamoja.pdf

USAID (2021). *Jifunze Uelewe ('Learn and understand')*. Usaid.Gov. August 18, 2021. Available from: https://www.usaid.gov/tanzania/documents/Jifunze-Uelewe

UWEZO (n.d.). Uwezo Na Jamii' community engagements for action. Available from: https://uwezotanzania.or.tz/community-engagements-for-action/ (accessed 20 April 2024).

"What's so Wrong with the Diffusion of Innovations Theory?" (2021). Available from: https://www.enablersofchange.com.au/whats-so-wrong-with-the-diffusion-of-innovations-theory/

Further reading

Asongu, S. A., & Nwachukwu, J. C. (2016). The mobile phone in the diffusion of knowledge for institutional quality in Sub-Saharan Africa. *World Development*, *86*, 133–47. doi: 10.1016/j.worlddev.2016.05.012.

Dooley, K. E. (1999). Towards a holistic model for the diffusion of educational technologies: An integrative review of educational innovation studies. *Educational Technology and Society*, *2*(4), 52–66.

Hung, D., Toh, Y., Jamaludin, A., & So, H. -J. (2017). Innovation becoming trajectories: Leveraging lateral and vertical moves for collaborative diffusion of twenty-first century learning practices. *Asia Pacific Journal of Education*, *37*(4), 582–600. doi: 10.1080/02188791.2017.1388213.

Marsh, D. R., Schroeder, D. G., Dearden, K. A., Sternin, J., & Sternin, M. (2004). The power of positive deviance. *BMJ : British Medical Journal*, *329*(7475), 1177–79. doi: 10.1136/bmj.329.7475.1177.

USAID (2019). USAID reading MATTERS conceptual framework | Global reading Network, November 2019. Available from: https://www.globalreadingnetwork.net/resources/usaid-reading-matters-conceptual-framework

USAID (2023). *Ten years of early grade reading programming: A retrospective*. Washington, DC: USAID. Available from: https://www.edu-links.org/resources/ten-years-early-grade-reading-programming-retrospective

Corresponding author
Simon King can be contacted at: sking72@gmail.com