79777

Effective Teacher Training in Low-Income Countries:

The Power of Observational Learning Research

Helen Abadzi[1]

Global Partnership for Education
November 1, 2012

GPE Working Paper Series on Learning No. 11

---

[1] Helen Abadzi is a Greek psychologist, who has worked since 1987 as an education specialist and senior evaluation officer in the World Bank and the Global Partnership for Education. She explores cognitive neuroscience applications that may improve the education of the poor. Her publications helped raise early-grade reading fluency to a high-level international priority. Email: habadzi@worldbank.org, habadzi@gmail.com.

Training and Performance Issues of Teachers in Low-Income Countries

The Education for All (EFA) initiative depends on students being taught by suitably and sufficiently trained teachers. But time-on-task studies conducted in low-income countries show that relatively little time is being spent on instruction, including the critical teaching of reading. Teachers may be absent often and may avoid teaching when in school (Abadzi 2007). They may engage with the few students who can do the work, neglecting the rest (Llambiri 2006, Abadzi and Llambiri 2011). They may fail to use textbooks even when they exist and spend class time copying on the blackboard. The same issues affect supervisors and principals (Abadzi 2006). As a result, students may graduate or drop out illiterate.

Donors and governments have invested a great deal in teacher training activities for lower-income countries. The investments in teacher training are potentially valuable, but need to be linked to results. Thus far evidence is limited. Preservice training often lasts 6-9 months compared to 3-4 years in higher-income countries and may be insufficient to remedy students' academic deficits or teach them how to teach. Textbooks on teaching are typically unavailable. (Distance education that is available in many countries provides limited texts for study with occasional attendance at meetings.) Inservice training typically involves lectures and some group work at training centers or, more recently at the school level. The cascade approach is used which may degrade the messages and result in wasted time and money. (e.g. IEG 2007). Lectures may be an efficient means of delivering content to well educated audiences, but they individuals must be able to elaborate information fast and take notes fast. People cannot easily translate theory or verbal instructions to procedures and decisions (Feldon 2007). So, teachers may attend and even pass tests, but their classroom behaviors seem uninfluenced by the training content. With teachers of limited education in particular, several studies have shown limited or no effects of training on student achievement.[2] (e.g. Unesco 1998).

The poor results have disappointed governments and donors (e.g. IEG evaluation of training, 2007). In various meetings many derogatory remarks are heard about teachers and their lack of motivation or accountability. The persistent teacher training problems worldwide make it imperative to seek new means for changing behaviors, particularly for poorly paid teachers with limited education. This must be done relatively quickly and efficiently so that teachers can impart basic skills to their students. Since feasible educational methods seem to have been exhausted, it is useful to look for solutions in the field that studies behavior. And there, much is found.

How likely are teachers of limited education to carry out a set of activities when no one is looking? This may depend on some little-used variables. If they do not feel comfortable with ability to do what they must, they may be absent often or may be absent from class when in school.

- Fluency. To teach well, teachers should not have to think too hard about responses and next steps. They need to execute effortlessly chains of procedures related to teaching particular topics.

---

[2] A graduate course is offered at the University of Massachusetts-Amherst to "probe the unique challenges of teacher education in developing countries and other low-resource contexts" and look for alternatives, "Many of these approaches grow out of the failure of attempts to transfer methods from high-resource environments."

They also need to develop automaticity and rapid reactions in math or spelling, to correct errors instantly.

-Sense of timing. Teachers must develop a sense of how long an activity should last; they should also have an overall timing sense for the year's curriculum. By the same token, those developing methods and curricula must have a sense of how many activities can be easily doable in one instructional hour.

- Encoding specificity and state dependent learning. Teachers would be less likely to remember when in their classrooms material taught in a training center;

- "Spaced learning", For consolidation, training material must be contemplated over a longer period of time than the usual 2-5 day workshops.

- Get sufficient intrinsic motivation from training. Often teachers are told to 'just do it'. But training should enhance desire to do the job. They should understand why it is worth carrying out certain procedures. Trainers must hear concerns, modify activities. Motivation research needs to be considered, particularly intrinsic motivation (see Gagne and Deci, 2005).

- Have proximal goals. (e.g. reading basics in 100 days). Proximal goal setting can develop self-efficacy and intrinsic motivation from previously low levels (Bandura & Schunk, 1981, Manderlink and Harackiewicz 1984). This implies that students' courses could have a shorter duration than the frequent nine-month spans (e.g. they can be modular).

Often teacher trainers lack sufficient knowledge on these issues and may just lecture. But training is needed in:
-Content
-Fluency of activities, knowledge (mental math)
-Specific methods, classroom management
- Motivational aspects

Of the above issues, this paper concentrates on the need for effortless execution of procedural chains and the application of little-known research.

Human and animal brains contain areas specializing in repeating various actions with intention and areas devoted to planning and future-oriented thinking (mirror neurons, Rizzolatti et al. 2002). When these mechanisms are used appropriately, behavior change is almost instantaneous (Dowrick 2006, 2011). Studies have begun exploring this area for student-teacher relationships (e.g. Zhou 2012). This research could also explain teachers' apparent unwillingness to teach.

Limited Education and Cognitive Overload

One overlooked explanation for teachers' limited behavior change despite training is cognitive overload. Too many students talking, too many things to do simultaneously may overwhelm them, particularly when they are less educated. This theory has received much attention in

educational situations of higher-income countries (e.g. by John Sweller 1991). Below is a brief description of the basics and prospects for applicability to teachers of lower-income countries.

The mind has limited computational resources for various tasks at any given moment. Human working memory capacity is limited, and people can only do a limited number of tasks simultaneously. To pass through working memory various tasks must be learned and "chunked'' to the point of automatic performance. By definition, people with limited education, such as people likely to become teachers in low-income countries, have automatized fewer skills than more educated people. They may be able to handle only 2-3 items at one time and keep in mind only very few students.

Even for better educated teachers in lower-income countries, the capacity for formulating and solving complex problems such as those presented in teaching real classrooms may be limited compared to some "ideal" model (Feldon 2007). Some teaching situations call for immediate rather than reflective responses that probably preclude rational processing of information in making an informed judgment or decision (Feldon 2007). Processing instructional responses consciously is slow and effortful. Teachers simply have no time to think too much about what they must do when faced with the complexity of a real classroom. When too much information is being given or too much must be done, the working memory is overloaded, people may not monitor what they do (van Gogh et al. 2010), and some tasks are not done. And if teachers expect to spend much effort or perceive that they may fail, they may avoid certain topics or the work itself.

Less well educated people may suffer more from cognitive overload. Since they have automatized fewer skills, they may know content less well and must think hard for answers. Yet they are expected to do a lot simultaneously: devise comprehension questions to ask while calculating students' math answers on the blackboard, try to block out the noise from the neighboring classroom while remembering monitor comprehension and not show gender bias.[3] These tasks may well exceed a teacher's processing capacity. Thus, some tasks may be rarely or never performed (such as remembering to question girls or querying silent students). Some literacy methods expect teachers to use multiple props, such as flash cards or booklets, and these are often found unused.

---

[3] Sweller (1999; Chandler & Sweller 1991, in Feldon 2007) distinguished three categories of cognitive load for teachers: *Intrinsic cognitive load* represents the burden to working memory inherent in the semantic content required for a particular task, i.e. the content knowledge to be taught in a lesson and the related pedagogy. *Extraneous load* represents unnecessary content that takes up space in working memory: a teacher's attention to a distraction or irrelevant events, such as a visitor or noise heard through the wall from a neighboring classroom. *Germane load* is the minimum level of cognitive load necessary for effective instruction (intrinsic load plus unavoidable extraneous load imposed by pertinent situational constraints). This includes awareness of students' prior knowledge of the topic, assessment of verbal and nonverbal cues indicating their level of comprehension, and monitoring students' level of attention to the lesson. Germane load is sometimes associated with motivation and interest. Any cognitive load that is not germane to the target task inherently deprives the teacher of cognitive resources that could benefit target performance. (see edutechwiki.unige.ch/en/Cognitive_load)

Teaching tasks must fit comfortably in working memory. Otherwise cognitive overload results and some will not be done

About 7 items for simple text

About 12 seconds

Cognitive networks

Long-term memory

35
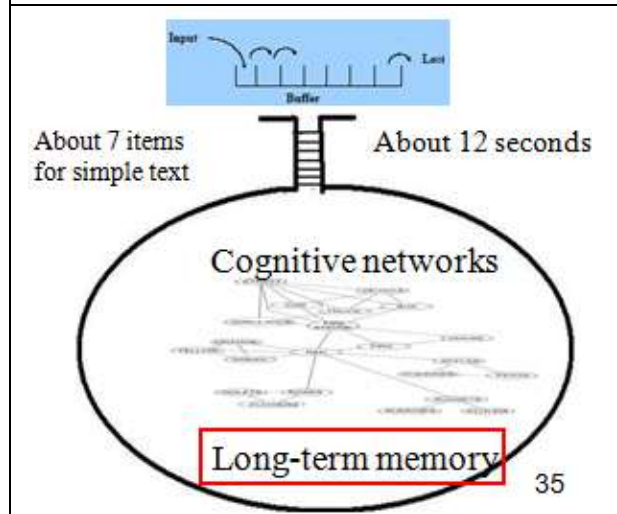
Cognitive load has been analyzed and measured in computerized environments but rarely in teacher performance. However, limitations in cognitive capacity may explain a number of inexplicable problems found in teaching and permit hypothesis testing. For example, "teacher centered" methods ought to create a higher cognitive load, but teachers have seen their own teachers do "teacher centered" methods extensively and may be effectively modeling that role. Furthermore, the tendency to just lecture may be an effort to minimize distractions while delivering content that may be complex. On the other hand, some teachers may prefer "constructivist" methods because students are supposed to take charge, and their own cognitive load may be reduced.

The cognitive load hypothesis may also offer a reason why some merit pay studies have had ambivalent outcomes (e.g. Adams et al. 2009). Merit pay is based on an underlying assumption that teachers know how to teach. (It also supposes that good teaching can be distinguished from poor teaching for purposes of accountability, that there is significant elasticity of supply, and that financial incentives will produce better results.) If the "elasticity of supply" is low, reliance on price mechanisms would be ineffective. Merit pay would have to be coupled with knowledge of effective techniques, so that the elasticity of supply can increase. Then prices can be an effective incentive.

There are potential solutions for keeping the cognitive load manageable and improving performance; an automatic pathway in our minds can generate fast and effortless nonconscious processing. The performance of *some* of those tasks (or *aspects* of all those tasks) must be automatized, and then rarely performed tasks may become more frequent. Teacher training methods can draw on cognitive approaches that emphasize the development of automaticity in teaching skills to minimize extraneous cognitive load and maximize effective performance. This document briefly presents a few possible avenues aimed at improving the efficiency of teacher and supervisor training. Such interventions are more complex than the default method of telling teachers and expecting them to retain and carry out the instructions. They are hypothesized to work because *the mind frees up resources with automaticity and may allow teachers to attend students better in class.* These include:
(a) video-based methods, audio-based methods such as interactive radio,
(b) goal-directed imaging, and
(c) intention to implement for enhancing prospective memory.

These techniques are not mutually exclusive. They can be combined and reinforce each other. In the sections below, the document describes the rationale and outcomes for these methods. Subsequently it proposes means for piloting and understanding better how these mechanisms work in the case of teacher training. Finally, various challenges and limitations are discussed.

A. Use of Video in Observational Learning and Self Modeling

The efficiency of audiovisual means for presenting material is broadly understood. They transmit complex interactions that would be harder to process by mere lecturing and thus may greatly speed up performance. The use of video for teacher training can be roughly of these types:
  (a) Informational, imparting content knowledge, such as science, teaching demonstrations
  (b) Feedback on performance, e.g. through the process called microteaching
  (c) Modeling others' behaviors in general terms; e.g. performance by master teachers
  (d) feedforward, often called positive self review (a modeling subcategory); feedforward refers to self model images of behaviors or skills never hitherto performed as such.[4]

Selectively imitating role models is a common human and animal process, which children do extensively. The process has been studied since the 1920s, and in the 1960s Albert Bandura experimented with the parameters of this phenomenon (Meharg and Woltersdork 1990). Some psychologists attempted to go beyond the usual phenomenon of playing videos of others for role modeling or for feedback. Instead they tried to improve athletic performance or modify the behavior of children with problems by filming various student actions and then editing them into videos which showed an ideal performance that had not yet been achieved. Thus, children saw themselves performing tasks that they normally did not. Since the 1990s, psychologists (Dowrick 1999, 2003; Buggey 2009) developed detailed procedures aimed at acquisition of certain skills. Individual videos are made that last 2-10 minutes and that depict people as able to carry out specific behavioral sequences. Their subjects watch them for 2-6 times during which they rapidly acquire skills/behaviors and achieve performance levels depicted therein. Some changes are quite amazing, and disabled children may become able to carry out movements previously thought impossible (video samples available). Even reading fluency can be increased through this method (Dowrick et al. 2006).

Dowrick and others describe this phenomenon as **feedforward and distinguish it from feedback.** Specific brain areas are aroused when someone is viewing a skill with the intention of repeating it.[5] In practical terms, feedforward is achieved either by taking a current skill and placing it in a new (challenging) context or by taking component behaviors and reconfiguring them into a significantly new skill (Dowrick, 1999, 2010). Rather than faithfully imitating, the subject is able to do 'mental time travel' and see oneself in the future performing a certain act. For this reason, when people imitate certain role models, they do not copy their performance exactly; they adapt it to their own circumstances. For example, teachers are widely known to reproduce the behaviors of their own teachers, particularly when they are less experienced. But

---

[4] Buggey (2009) uses "feedforward" for behaviors carried out less than 50% of the time and forward self review for behaviors carried out more frequently.

[5] When a person watches someone else perform a task with the intention of later replicating the observed performance, motor areas of the brain are activated in a fashion similar to that with accompanies actual movement. the amount of activity occurring in the intraparietal sulcus -- when watching to learn accurately -- predicts how well these actions are reproduced minutes later (Frey and Gerry 2006). Similar findings are also in Rizzolatti & Craighero 2004; Rizzolatti et al. 2006.

though they may reproduce the prior teachers' intent and activities, the behaviors are not automatized chains; they are adapted to the current settings.

The feedforward technique illustrates "successes not yet achieved" (Dowrick, Tallman, & Connor, 2005, 131). The theory of feedforward proposes that "component behaviors (in the repertoire) are reconfigured to produce a new skill or level of performance" (Dowrick, 2011, p.2). In contrast to *feedback*, which focuses on past or present performance, feedforward is a depiction of future performance. Video self modeling (VSM) projects can incorporate "feedforward", or images of future mastery (Dowrick, Kim-Rupnow, & Power, 2006).

Feedforward techniques are used to create video videos in which individuals can watch models of themselves learning new skills, developing their self-image of being capable of doing skills that they are not currently capable of doing independently. By starring in one's own video, the individual has a compelling and authentic model of future success – the 'self' (Graetz, Mastropieri, & Scruggs, 2005). A VSM project that incorporates feedforward depicts a skill or behavior not yet mastered, one that an individual might be trying to achieve more consistently or regularly. When filming the video, the individual can be provided with support or a scaffold to reach mastery. While editing the video, the scaffold is edited out in order for the individual to be depicted performing the skill independently in the VSM video. Eventually the need for the scaffold to perform the behavior is eliminated. (It is unclear how this would apply to teacher training.)

For example, in one scenario the tutor provides a scaffold, by modeling fluent reading. During the echo reading process, the student demonstrates fluent and error free reading, with the assistance of this scaffold. While editing the video, the tutor cut out the portions in which she can be heard reading, leaving only the video in which the student is reading the passage independently and fluently. The edited video was a depiction of the student as a fluent reader, providing images of future mastery of the target skill.

Feedforward may be the fastest way to learn such behaviors; it can be twice or sometimes 10 or 100 times as fast as other methods (Dowrick 2011). Feedforward can provide examples of best or better ways to teach, excluding as much else as possible. Such clear, uncluttered information can be absorbed instantly, for potential reproduction in practice or in life. Additional information may cause cognitive overload and slows down or prevents learning.

For feedforward to work, it must include only skills that are component behaviors in the repertoire of the observer. For example, new teachers should be able to hold a notebook with scripted lessons on their arm, speak clearly, remember and say the words and sounds, listen and discriminate. The video puts all these pieces together in the right configuration. The observer puts the image of him/herself into his/her brain - as an image of future behavior.

There is recent neuropsychological evidence for this view, in mental time travel, including neuroimaging studies (e.g., Addis, Wong, & Schacter, 2007) and in mirror neuron studies (e.g., Fogassi & Ferrari, 2010). There is clearly increasing evidence and opinion in support of feedfoward strategies from independent sources (e.g., a "feedforward interview" for job applications or periodic reviews; aspiring marathon runners are advised to create two "mental tapes" of 1.5 minutes for last best performance and for the way they want to see themselves run; Whitsett et al. 1998).

Forward self review: The most rapid behavior change can occur from seeing oneself be successful in challenging situations (Dowrick, personal communication)

In teacher training, modeling has been used in various situations intentionally and unintentionally (e.g. when teachers imitate their own). However, *feedforward* has hardly been used. The research suggests that teachers could be helped to acquire, perform, and become fluent in various behaviors that they normally would not carry out (Table 1).[6] Videotaped self-modeling has been used in individual cases to allow subjects to view themselves as they might behave in the future. Thus, it could be sustainable in high-income countries, where teachers are personally monitored and extensively mentored. In lower-income countries teacher training takes place on a large scale, and it is not possible to monitor individual teachers or make individual tapes.

To develop such feedforward and in general imitation of movements for teacher training purposes research would also suggest stopping in certain places to create chunked components (Agam and Sekuler 2008). When analyzing seen actions, the brain privileges regular, consistent curvatures, as well as grouping components that form a coherent shape into a unified "chunk." Inconsistencies among the directional components of a motion sequence cause the sequence to be chunked into additional components, which increases the load on working memory, undermining the fidelity with which the sequence can be imitated.

The viability of this powerful neuropsychological system should be studied in larger-scale applications of lower-income countries. Mere modeling in the US by master teachers may have limited effects (e.g. Joyce and Showers, 2002). Would peer videos be effective in countries where individual monitoring is feasible? Some examples exist though they are hypothesized not to work as well (Buggey 2009). The mental time travel hypothesis suggests that models would have to be relevant, that is similar to those that a teacher would be likely to see as him/herself or already successful. Also, the more similar somebody is to the model s/he is watching the more

---

[6] For special education students at least, acquisition deficits are either a lack of knowledge on use of a skill, even under optimum conditions, or failure to know which skill to apply in specific situations. Performance deficits occur when a person has learned the skill but fails to apply it at acceptable levels in appropriate situations. A third type of social skill deficit is a fluency deficit, which is characterized by knowing the skill, wanting to use it, but lacking the ability to demonstrate its "polished" performance. Fluency deficits require repeated practice with feedback in order to fine-tune the social skill (Gresham et al. 2001).

likely it is that the person will pay attention to the model (Bandura 1969). More recent research, however, suggests that people are more likely to imitate the behaviors of someone with whom they share the same goal (Ondobaka et. al 2011).

For example, teachers could be exposed to brief videos of a closely matched model carrying out specific behaviors that are rarely observed, such as monitoring student work during individual seatwork. Some could model before other teachers for microteaching purposes if there is time in a training event, or the trainees could carry out specific sequences in unison during the event if this is appropriate. The expectation would be that the teacher would return to the classroom and carry out this performance. Thus the teachers would be told that they must watch with the intent to repeat later, because watching with the intent to repeat later has specific effects in the brain (Frey and Gerry 2006).

In the case of special education, self-modeling videotapes are typically two or three minutes in length. Critical tasks are defined for videotaping. A typical schedule for viewing is 6 to 10 viewings within a two week period, either at home or at school. Sometimes the tapes are reviewed again after two or three months when a maintenance "booster" is desirable. These frequencies could be used as a guide for planning teacher training. (An pilot preparation of videoclips has taken place in Liberia in March 2011).

However, the video methods by themselves would not be sufficient in changing behaviors for the long term. For successful forward self modeling, teachers must know why certain behaviors are needed (Dowrick 2009). Thus, they must understand the rationale for doing the various displayed activities. On the other hand, evidence from children suggests that mere discussion of solving problems has weak effects that are not maintained over time (Gresham et al. 2001). Prospective planning must be an integral part of the training (see sections below).

If this teacher watches an unedited version of his session, it will be feedback (or microteaching). If he watches only the best, error-free segments, it will be self modeling, known as positive self-review, which strengthens erratic, newly learned skills.

Economists may ask what **incentives** teachers would require to do this. They are often underpaid, and no one will watch them to know they did the right thing. The observational learning theory leads to the hypothesis that since people are biologically set up to model actions of similar people, these exercises may lead to behavior change in the class with limited effort. Furthermore, people seem to be set up through evolution towards social learning and conformity (Perrault, Moya, & Boyd, 2011).   The question is the magnitude of effect in the short and the long term. After all, the brain has an extensive reward mechanism, and every action is in some way affected by it. Months down the road, the teachers must see reasons for

carrying out certain behaviors; if they learn incompatible ones, the earlier ones may become less frequent.

Alternatives to feedforward may be audio modeling through interactive radio (see Annex).

B. Goal-Directed Imagery – Automatizing Actions that Require Effort

As with forward self review, visualization of actions as completed seems to facilitate their execution at a subsequent time (Liu and Park 2004). People imagine themselves as successful and achieve imaginal self-modeling. Goal-directed imagery stimulates the parts of the brain responsible for actual sensation, and images that can be named easily are better remembered. It may even increase muscle strength.[7] Furthermore, goal imagery seems to enhance implicit motivation (Schultheiss and Brunstein 1999). Thus tasks ought to be more easily recalled at specific times of the day if people visualize in advance doing those tasks.

Several applications of this technique can be found in health and sports psychology. Visualization studies suggest that the technique helps dieters stick to healthy eating, assists women in remembering to go for breast examinations, and increases the use of medic-alert bracelets by those who need them. The young and middle aged can also benefit from visualization. This method may be strengthen the effect of visual experiences, because weak effects are noted in the literature (Dowrick, personal communication).

For example, 31 people over 60 were trained to track their blood sugar several times a day using a standard testing device, much as diabetics must do. Study participants were put into three groups: a **deliberation group** talked over the reasons why daily blood sugar testing was a good idea; a **rehearsal group** recited the instructions for using the testing device; while an **imagination group** spent three minutes imagining themselves using the glucose monitor within the home or work environment. Results showed that the group visualizing an event was much more likely to remember to do it when they returned to their homes (Liu and Park 2004).

This methodology could be used not only to automatize complex behaviors but also to overcome the "state dependent learning" effect. This memory feature encodes learning on the basis of location, so teachers who may learn certain skills in a training center may forget them inside the class. For example, recall is better in the same room (Smith, Glenberg, and Bjork 1978). But if participants thought about the room in which the training had taken place, there was no difference in performance (Smith, 1979; Reisberg p. 168). Thus visualizing the intended actions taking place in a class would increase the probability that they would be done.

C. Intention to Implement – Enhancing Prospective Memory to Execute in the Future

The visualization technique shown to be effective is greatly enhanced by planning. Forming detailed implementation intentions can increase the probability that a future behavior is actually

---

[7] Wright and Smith (2009) used self-modeling videos to "prime" a period of reflective imaging of workouts doing arm curls. They found that 30 minutes of curls plus 30 minutes of imaging produced slightly more muscle mass than full 60 minute workouts.

completed. Mere rehearsal of the procedures and intentions to complete the action were less effective than planning how to carry out the required tasks (Liu and Park 2004). This is particularly important for tasks that people must initiate on their own, such as teaching activities. There are neurological mechanisms that account for such activities (Jeannerod 1994), including mental time travel type prospection (Gilbert & Wilson, 2007).

Examples from health psychology show again that people are more likely to carry out future intended actions if they plan them in detail (Kliegel et al. 2007). A meta-analysis of 84 studies revealed medium-sized effect sizes of implementation intention interventions on goal attainment (Gollwitzer and Sheeran 2006).

For example, among smokers, the cessation percentage increased by about 30% (from 10.3% ot 13.4%) if they planned with interviewers how they were going to deal with quitting (van Osch et al. 2008). Similarly orthopedic patients persisted with exercises longer if they planned ahead what they would do (Ziegelman et al. 2005, 2007). In other experiments, participants who formed an implementation intention were more than twice as likely to self-initiate the intended behavior (e.g. writing down the day of the week on every sheet of paper received during the experiment) compared with participants who either were merely instructed to do so or actively rehearsed the instruction (Chasteen et al. 2001). Similarly, implementation intentions alone significantly increased consumptions; but targeted mental imagery of the cues and cue-response link doubled the consumption of fruit among low consumers (Knäuper et al. 2011).

Another study tested a *"directed thinking"* paradigm. College students who did not exercise on a regular basis or exercised inconsistently were asked to think about ideas that fell into either the "reasons" category or the "actions" category. For example, some participants were asked to list the reasons why they should increase the performance of a target cardiovascular exercise they had previously selected, such as to be healthier or lose weight. Other participants were asked to list **actions** they could take to increase exercise performance, such a joining a gym or working out with a friend. During 8 weeks, bringing to mind actions to increase exercise performance led to an increase in exercise and improved cardiovascular fitness. By contrast, students who repeatedly just brought to mind the reasons why they should do the target exercise did not increase time spent exercising (Ten Eyck, Gresky, & Lord (2007).

To improve prospective memory, the plan should specify the concrete situation that is appropriate for initiating an intended action and link that situational cue to the intended action. However, teachers should be relatively fluent in the tasks that they otherwise do to carry out their intentions. If tasks are cognitively very demanding people may forget to do them (McDaniel and Scullin 2010). This research suggests that teachers in training must discuss among themselves or with others details about how they would carry out specific behaviors in classroom following training. This would add benefit to any training event, even those involving lectures.

Implications for the Training of Educators in Lower-Income Countries

 Teacher education programs focus on pedagogical theory and conceptual frameworks in hopes of instilling flexible knowledge that can be adapted to unique classroom situations. But even the

US student teachers typically find it hard to convert theory into classroom practice. Unless programs offer extensive and well-scaffolded practice, skill automaticity does not develop (Fisk 1989).

The inability of teachers in low-income countries to learn teaching with just a few weeks' preparation given the usual methods is not surprising. It typically takes years of experience to automatize classroom routines, and the ones automatized by such teachers may produce poor learning outcomes. If a number of needed routines were established through observational learning and intention encoding methods, then teachers might use textbooks more effectively and reduce the incidence of extensive copying. The development of automaticity is particularly important for feedback and monitoring of students. For example, teachers may have more cognitive resources to think and calculate the right answers of math problems students have written.

On the basis of the above it is hypothesized that teacher training efficiency and efficacy might substantially increase with a combination of three domains:

- Actual content knowledge, i.e. geography
- Automaticity in certain content areas, e.g. training in mental math to spot errors
- Subject matter methodology and classroom management

For methodology and classroom management, observational learning has considerable potential. The techniques to be used would be:

(a) informational videos to explain the rationale and teach content,
(b) peer-oriented feedforward videoclips to model specific desired behaviors[8]
(c) visualization activities to facilitate execution of desirable behaviors at a future time, generalize situational cues to overcome state-dependent learning,
(d) immediate practice in a group mode if possible, to maximize time on task
(e) prospective planning activities, asking students to contemplate how they will carry out certain behaviors in their classrooms

The videos would aim to increase the incidence of certain behaviors such as (a) common activities to carry out in order to teach reading, such as the common method of "I do, we do, you do", (b) classroom management to direct attention and feedback to lower-performing students who often drop out illiterate, (c) motivational messages about the goals of Education for All, discussions of experiences with childhood teachers, diagnostic insights about student achievement.

The priority subject matter is reading instruction. Results of distance-based training and supervision of Headstart teachers in the US (Powel et al. 2010) suggest that teachers can learn the relevant skills by video, and pilot initiatives in Liberia and Egypt are also promising. Aside from reading, teachers have limited knowledge of math and cannot do many simple calculations

---

[8] Because of the role modeling potential, it may be less useful to show teachers their errors on video; for informational purposes, however, this would be useful in preservice teacher training.

automatically. Perhaps methods based on observational learning will help with math.[9] They may also increase automaticity in the higher content levels, though this topic is beyond the purposes of this report.

The above means can also help train supervisors, principals, district directors and higher-level administrators). Currently a training method predicated by some donor agencies is "communities of practice"[10] and such materials would enhance its effectiveness. The informational aspects of video can show them the classroom patterns that are linked to positive and negative outcome and offer them fluency in recognizing the various patterns. Also, they could watch models of supervisors carrying out their work as positive self-review. These would enhance the probability that they will actually go to class and carry out the expected work tasks. Fluency and better discrimination of correct and incorrect patterns may reduce the work avoidance observed among these.[11]

Academics have been working on the details of obtaining such video.[12] It is unclear at this point how the technique must be specified in the case of training videos that are not personal. Some important characteristics of a feedforward video (Anderson and Dowrick 2000) are:

- Show only successful and independent performance of a skill beyond one's current ability;
- Show only positive images. Avoid correcting errors
- Focus on one attainable skill; if the skill is complex, consider making a series of videos that demonstrate the sequence, step-by-step
-Keep the video short. The most successful feedforward videos are about 2-3 minutes long.

Technology is becoming increasingly available and adaptable to poorer areas. Many cell phones these days have video capability. Pico projectors can project to groups of teachers from a cellphone to the wall. Thus videotaping can be done by amateurs who know the subject rather than videographers. It is not important to have high-quality video, the picture quality merely has to be satisfactory. USB cameras, such as the Flip and the Sony Webbie, cost less than $200 in the U.S. There is a need to know how to upload video, edit it, choose needed segments, integrate them with training activities. Sustainably usable technology is needed, particularly for carrying digital content to areas with sporadic or with no electricity. Inexpensive laptops can be connected to small, battery-powered projectors (e.g. Optoma brand that costs around US$300).[13] Extra

[9] Elementary math content and exercises can be potentially delivered through cell phones or low-cost computers in the countries that have them.

[10] A publication about creating teachers' communities of practice laments the traditional teaching methods and imitation of teachers' teachers and encourages them to " ...think for themselves and go beyond the recall of facts and mimicking of skills" (du Plessis and Muzaffer 2010).

[11] Interventions with supervisors could be piloted by governments in rapidly expanding NGOs (e.g. IEP in Mali). To understand specific aspects of their work that are not satisfactorily carried out, supervisors could be "followed" for a few days with a remote camera that they would wear. Supervisors would also receive clear behavioral checklists with a few critical variables. These could also be given to teachers during training. IEP in Mali has developed such a supervision checklist that could be revised or adapted according to content areas.

[12] **http://www2.ku.edu/~kucrl/cgi-bin/drupal/?q=node/225** (University of Kansas, Lawrence 66045 (785) 864-0693, Shawn Smith, seanj@ku.edu. http://www2.ku.edu/~kucrl/cgi-bin/drupal/?q=node/221 Do it Yourself: Video self modeling made easy: https://sites.google.com/site/diyvsmreviewsite/

[13] Using smartphones equipped with a system allowing internet access through cellular network, teachers accessed lesson plans and content in science and language arts: http://www.learningandleading-digital.com/learning_leading/20100304#pg18" (Burns et al. 2010). Also sending self modeling prompts to cell

batteries that can be carried to areas where there is no electricity and can be charged from cars. Solar chargers of computers could be used for longer-term stay.

Evaluation. Ultimately, behavior changes should be reflected in improved learning outcomes. Thus effectiveness must be evaluated in comparison to baseline information of teachers' classroom behaviors. Also attitudes and beliefs about teaching efficacy given the above techniques ought to be measured through behavioral instruments developed to track changes. To establish and test a causal chain student performance data would be needed.

Experimentation must also determine but whether the cost and complexity are worth the effect size. For some methods the cost-benefit analysis may be easier than others. For example, planning discussions on how to implement a change or visualization exercises can cost very little and (if done correctly) can be a training activity without large-scale decisions.

Methodological Features for Development and Piloting

There is a need to pilot likely procedures and develop a specific methodology and toolkit. For the sake of sustainability, this work must be done in close collaboration with governments or influential NGOs.

The essentials of developing vidoclips would be:

- Identify the target skill or behavior (e.g. reading fluency, how to share)
- Develop a plan for filming a short 2-3 minute video; the plan can be a storyboard or script in some cases
- Determine what the "scaffolds" are (coaching provided by the teacher/tutor)
- Video-record the trainee performing the target skill or behavior with scaffold; for some behavioral interventions, such as a video on "how to share," it is appropriate to involve peers in the video in order for the student to be in an authentic interaction
- Edit the video, cutting out the portions with the scaffold
- Have the trainee view the edited video in which the target behavior is being successfully performed
- The training videos would be stopped for discussions and would be replayed multiple times (see Pratham example in Annex).

For reading in particular but also perhaps for early-grade math, a routine ought to be established that teachers would follow every day.[14] Thus teachers would receive:

---

phones has been successful in USA (Dowrick, personal communication).   Where GPRS is available, perhaps podcasts are possible.
[14] Specifically for reading, scripted lessons ought to be videotaped, as the Research Triangle Institute has done in Liberia. It must be understood better whether teachers need to watch every one of the 90 or so lessons developed for Liberia in order to carry them out fluently, or whether they can become fluent in carrying out component behaviors and whatever script they receive. In principle teachers could watch the lessons through cellphone formats already developed by the Educational Development Center in 19 schools of Mali. They could also receive the content in mp3 players, potentially to use like interactive radio.

- A video of 10-15 minutes duration showing a demonstration of every piece of the routine
- Explanatory videos of 5-10 minutes to illustrate specific techniques and their rationale
- Feedforward videos within each of the segments in the routine

Research questions to ask would include:

- What type of model would the teachers in a particular country be more likely to imitate, envisage themselves in the future? (distance in terms of SES, gender, age, accent etc)
- Dosage of treatment: What duration of exposure to the video, in what interval to repeat?
- How many discrete behaviors can be modeled in a single teacher training session? The importance of distributive learning as well as the primacy-recency effects must be understood for visual stimuli.
- How can audiovisual training sessions be delivered through cascade training? Will it be possible to create a pyramid of master trainers? Would a DVD or mp3 audio function like interactive radio with respect to training of trainers?
- What are the magnitudes of effect that can be obtained through this method?
- Does projection screen size affect the adoption of desirable behaviors?
- Utility of discussing teachers' memories of their own (problematic) teachers, exploring status issues in their own classes. (Teachers were likely survivors of massive dropout.)
- Utility of explaining to educators the imperative of teaching all students.

To understand these issues better, a "field laboratory" could be established, that is a place to create video-based protocols informally, try them on willing groups of teachers, obtain feedback, and reconsider before scaling up to a pilot. Behavioral instruments should be developed to track changes. Countries or teacher training programs are sought out to play that role. Consultants could travel to that location to experiment.

At least two months would be needed to develop content and videotape teachers in specific countries who would carry out the various activities. Costs would have to include the suitable electronic equipment.

After pilots, training ought to start through supervisors and inspectors, who can then lead about it. School-based and center-based courses of various duration would be tried. The videos would be offered in preservice as well as in inservice training. An organized and easily comprehensible document would be developed for translation in multiple languages and for use by trainers.

After formal or informal trials, workshops can take place. A technical panel can also be convened to discuss findings and propose further refinements. The panel would consist of academics who have used feedforward techniques (mainly special ed.) as well as teacher trainers of low-income countries.

A toolkit containing very clear and detailed instructions (including videos) would be developed and disseminated to international agencies and governments in multiple languages.

The timeframe constraints need to be considered. An alternative approach to professional development would focus on clear instructional goals for all teachers, and measure success

against progress toward these goals. That is, teacher training should be linked to assessment in real time. In principle, it should be up to the teachers and teacher trainers to find approaches for which there is support in the literature, and then see which of those would achieve results for children. In reality, information about these and other methods may not be available to teachers and trainers at local levels of low-income countries. Hopefully progress will be made in this respect.

Type of teaching improvement needed	Teacher behaviors	Likely effective intervention type for teachers	Comments
Content knowledge	Does not know the material needed to teach	Classroom and distance content	May be developed more generically or obtained from other countries if practical
Basic skills automaticity	Cannot do various tasks easily, such as calculate quickly	Practice in calculating to automaticity	Practice skills such as mental math in various group settings, possible use cell phone to send quizzes to teachers
Behavior acquisition	Can't do a skill Lacking in repertoire Missing, confusing specific steps in performing sequence of teaching skill steps	Modeling Coaching Performance feedback from other teachers and principal	Brief videoclips with the specific skills, repeated, considered Microteaching events if possible Visits by coaches (who must be similarly trained)
Performance	Won't do skill Has skills but fails to use at acceptable levels in given situations	Increase motivation Manipulate antecedents and consequences	Forward planning, consideration of reasons Videoclips with specific skills
Fluency of a behavior	Need to fine tune skill Has skill and wants to use it	More practice Behavioral rehearsal Differential corrective feedback Reinforcement for fluent performance	Forward self-review if feasible Videoclips with specific skills if impossible to self-model
Absenteeism, time wastage	May not really know what to do in class	A special course could be developed for teachers with absenteeism issues	Get teachers to identify their reasons for poor performance; give all the above

Annex

Modeling Audio or Video Instructions

To counteract the dearth of trained teachers in low-income areas some initiatives have focused on providing live lessons by directly to students. These must be mediated by teachers or community workers. As the teachers get involved with these, they gradually acquire the knowledge that is transmitted and model the behaviors. Two examples are interactive radio and Indian Digital Study Hall.

Interactive Radio. Interactive radio has been used in many countries and financed by the Bank and USAID since the 1970s with positive results for students. But it also constitutes daily teacher training. An Educational Development Center (EDC) interactive radio evaluation showed teachers over a short time demonstrating a better understanding of the pedagogical concepts emphasized by broadcasts and utilizing them.[15] So, interactive radio instruction seems to provide detailed audio-based role modeling. Since teachers are on location. They can enact what is learned and practice it quickly. Interactive radio may also increase time on task and serve an accountability function.

Any radio will receive the transmissions, but USAID has been distributing this model that provides other benefits as well.

Interactive radio exists in many countries, notably the Democratic Republic of the Congo, Guinea, Guyana, Haiti, India, Indonesia, Madagascar, Malawi, Mali, Nigeria, Pakistan, Somalia, Sudan, Tanzania and Zambia. However, its coverage is not universal, and its startup costs have been very high. For example, it costs about US$1 million to develop content for one grade per year. Carrying this out for the long term in multiple grades implies somewhat complex logistics. Also transmission clarity may vary from day to day. So for a number of reasons interactive radio is gradually getting abandoned.

One would expect that technological improvements would result in new means to distribute content and teacher training, but this has not yet happened. The transmissions fit into mp3 players (see picture) and are given to schools in some areas, but it is uncertain whether they are used. The specific hour of the transmission functions as a deadline for teacher and student attendance, and if it is optional, it is unclear whether it will be done.[16]

Means must be found to maintain the daily modeling effect and perhaps deadline aspects of interactive radio. The mobile network can be used perhaps for radio transmission of interactive radio materials. Visual content would be ideal, but the equipment and knowledge needed are more complex than a transistor radio.

---

[15] Ho and Thukral 2009. There are many evaluations and reviews of this medium.
https://secure.edc.org/publications/prodview.asp?2017
[16] The Pt. England school in Auckland, mostly Pasifika children have effectively used podcasting, interviews amongst themselves, book reviews, 9-10 yr olds, overall significant contribution to literacy – tens of thousands tuned in all over the world, including non-English speaking countries (Dowrick, personal communication)

Prospects for replacing the teacher role: Digital Study Hall

Since 2005 an Indian organization has been engaged in the creation and distribution of instructional videos. [17] The project is a collaboration between computer scientists and education experts. The main aspects of DSH are: A "people's database of everything", a network of hubs and spokes to distribute videos to schools, and technology for sharing community-generated video.

The Digital StudyHall seeks to improve education for the poor children in slum and rural schools in India. They digitally record live classes by the best grassroots teachers, transmit them through the postal system, collect them in a large distributed database, and distribute them on DVDs to poor rural and slum schools. The principal means of disseminating the content in the DSH database is shipping DVDs to spoke schools. Each spoke school is given at least a TV and a DVD player. Many schools also need a big lead-acid battery and a DC-to-AC inverter for dealing with intermittent electricity.

Educators use the system to explore pedagogical approaches involving local teachers actively "mediating" the video lessons. By harvesting a "viral phenomenon" of community participation, the organization aims to help train teachers and deliver quality instruction to underprivileged children. They run pilot "hubs" in Lucknow, Calcutta, Pune, and Dhaka, covering more than 30 schools. They had collected more than 1500 recordings of lessons in English, math, and science, in Hindi, Bengali, Kannada, Marathi, Tamil, and English, and 1500 additional videos of other materials such as stories, special science and history topics, and training sessions. They also started applying the same approach to agriculture extension work (Digital Green) and awareness campaign for rural healthcare (Digital Polyclinic, operating in Lucknow and Ghana).

The project refers to the process as "mediation-based pedagogy". A teacher (or a "mediator") is placed in between the students and the TV. The mediator periodically pauses the video and engages the students in various activities based on what has just occurred on TV. These activities may include asking questions, inviting kids to do board work, and organizing role-playing activities. The mediator's job is to make his or her class as lively, dynamic, and interactive as the one conducted by the model teacher on TV. In effect, the video and the mediator form a "team:" the video provides an example, a framework, a lesson plan, and a content and methodology model; while the mediator, who may not be highly skilled in some domain-specific knowledge, supplies the crucial interactive element.

According to the organization's website, highly motivated teachers can study the supplied videos on their own, ahead of the live classes. In traditional teacher training workshops that last just a few days, the short duration necessitates that the topics covered must be kept at an abstract level, and it is not always clear how such abstract principles should relate to many of the daily topics. In DSH, the videos carried home by the participating teachers provide an ongoing and highly specific training: the local teachers learn by observing the best model teachers in action day-in and day-out, and they learn by doing. So this mode of training has the potential of being much more effective if teachers are willing to put in the time and effort.

---

[17] http://dsh.cs.washington.edu/info/descr_pedagogy.html

Teachers using recorded lessons on TV

Another variation of the theme is "peer-mediation," the approach of recruiting the best-performing fellow students to serve as mediators during periods when the local teachers are absent, which are common occurrences in government schools in India. The website of the organization reports that student mediators appear to display a high degree of responsibility and enthusiasm when they are put in charge.

Students conducting classes mediated by TV | **Source: website**

Clearly this approach works in environments where there is considerable technology available for the better off. An evaluation presented at an American Educational Research Association conference showed a dramatic rise in student test scores of participating schools, significant improvement of subject matter knowledge and pedagogical skills of local teachers, and increased student participation. The Digital Study Hall won the 2007 ACM Eugene Lawler Award for Humanitarian Contributions within Computer Science and Informatics, and the top prize in the education category of the 2008 Tech Awards by the Tech Museum of Innovation.

Example of Extensive Video Use by the Pratham NGO of India

The Indian NGO Pratham is extensively using video for training in language and arithmetic, along with a textbook. Pratham teaching-learning activity package in Read India this year is called CAMaL (Combined Activities for Maximized Learning). The fundamental element of

CAMaL is to combine activities: listen, say, do, understand, read, write, solve. Many videoclips have been uploaded to youtube.com.[18]

Pratham found out that it was helpful to know the audience well. They showed a video fully first and then did a second round where the trainers stopped and commented or discussed what was seen the video (with frequent pauses). Pratham trainings are often accompanied by several days of "practice" classes (rather than mere "demo" classes), juxtaposing what the teacher wa doing in the video versus what teachers themselves were able to do in their practice class. Once the teaches saw the videos they wanted copies to watch on their own, so Pratham distributed them as well.

Suggested Procedures and Topics for Feedforward Modeling

Professor Dowrick proposed the following process in countries or environments that may be able to carry it out:

- Hand pick about 20 promising teachers; select across age groups, genders, physical features, language/accent;
- Establish (with local collaboration) about 20 teaching issues, strategies, behaviors (called "teaching content") to be portrayed audio-visually and also to be described in words.
- Work with teachers to create video examples, all Teachers in all content areas;
- Involve teachers to the max, where possible storyboarding, video recording, editing under the experts' supervision.
- The teachers become trainers of next cohort; use videos above, create more videos and training materials;
- "Teach one thing, and teach it well": limit the number of teaching content examples to the ones that are most important; work very hard with first cohort to ensure that they become successful teachers and trainers of other teachers.
- Expand exponentially until whole country is trained. (In 6 months, if every two teachers could train 20 more, that would be 220. In another 6 months, there could be 2,420 trained teachers, and potentially in three years there could be 35,431,220 (!!)
- After 6 months start phase 2: Add new content area (for example, in-class peer tutoring, where the better third of class achieving words per minute criterion) would tutor one or two students in the rest for 30 minutes once a day using ACE/PALS combination of methods.(?)

---

[18] Pratham made the clips below at a public primary school (in Dadri, Gautam Buddh Nagar district, western UP)
Clip 1 : Mind mapping 1: http://www.youtube.com/watch?v=MoKesJnzN0g [may have moved]
Clip 2: Mind mapping 2 :http://www.youtube.com/watch?v=zN_afW63uiw [may have moved]
Clip 3 : Working on punctuation :http://www.youtube.com/watch?v=8jebuO_gSSo
Clip 4: Textbook- related activities by a teacher (panchayat shikshak) in a classroom in West Champaran district in Bihar http://www.youtube.com/watch?v=LO-tqoxF3v8
Angrezi A to Z about teaching English to primary school children in India by Pratham teams in Solan, Himachal Pradesh. After training, government school teachers successfully implemented the pilot project in their schools (http://youtube.com/watch?v=r2SiEMj3SGc)
Another 8 training videos on teaching English and around 6 videos on teaching Math
http://www.youtube.com/user/prathamorg

- In year 2, start phase 3: in Year 2 and every subsequent year, repeat Phases 1 & 2 until every teacher is trained and every student is literate.
- To the extent possible, do on-the-job training and train-in-place (not train-and-place)
- Follow up and follow through to maintain integrity of training; update and revise (but not expand) teaching content areas.
- Start where video is easily available

For example, a problem behavior worldwide is teachers' tendency to "broadcast" to all students and be satisfied with the participation of a few voices, often ignoring most of the class. (This problem leads to eventual dropout.) For this and other problem behaviors, videos would model behaviors such as below. (List is subject to revision by expert and local panels.)

(a) Querying techniques, such as calling on all students at random; naming a student first and then asking the question;
(b) Offers of quick and corrective feedback; explaining why a child made a mistake rather than ask the same question over and over
(c) Providing intrinsic incentives to all students, asking all to reply
(d) Showing how to structure information vs. comprehension questions
(e) Moving away from the front desk position to attend to the back
(f) Monitoring students' seatwork rather than sit unengaged and wait for them to finish
(g) Ensuring that students can sound out letters and numbers, can actually decode text rather than merely repeat what they heard
(h) Monitoring individual students momentarily, e.g. listening to every child read for a minute a day while others do individual work (example exists)
(i) Using in application segments of sentences that students recite
(j) Giving and controlling elaboration exercises ("active learning")
(k) Use of textbooks in various courses
(l) Writing clearly on the blackboard in grades 1-2 so that children can read the handwriting
(m) Using simple aids such as pocket charts and pre-written letters and words
(n) Making culturally appropriate gestures of encouragement and praise
(o) Learning about and teaching morphology, how to break words down (e.g. Deacon and Kirby 2010).

References

Abadzi, Helen. 2006. Efficient Learning for the Poor: Insights from the Frontier of Cognitive Neuroscience'. World Bank: Directions in Development.

Abadzi, Helen. 2007. Instructional Time Loss and Local-Level Governance. *Prospects*, 37(1), 3-16.

Abadzi, Helen. 2007. Absenteeism and Beyond: Instructional Time Loss and Consequences. World Bank Policy Research Working Paper Series no. WPS4376

Abadzi, H. and S. Llambiri. 2011. Selective teacher attention in lower-income countries: A phenomenon linked to dropout and illiteracy? *Prospects*, 41:491-506

Adams, S. J. Heywood, J. S., and Rothstein, R. 2009. Teachers, Performance Pay, and Accountability—What Education Should Learn From Other Sectors. Economic Policy Institute.

Addis, D. R., Wong, A. T., & Schacter, D. L. (2007). Remembering the past and imagining the future: Common and distinct neural substrates during event construction and elaboration. *Neuropsychologica*, 45, 1363-1377.

Agam Y, Sekuler R..(2008) Geometric structure and chunking in reproduction of motion sequences. *Journal of Vision,* 8(1):11.1-12

Anderson, J., & Dowrick, P. W. (2000). ACE reading video enhanced tutoring: Feedforward guide. Honolulu, HI: Center on Disability Studies.

Bandura, A. 1969. *Principles of Behavior Modification*. New York: Holt, Rinehart & Winston.

Burns, Mary, Marianne Montalvo, and Rebecca Rhodes. 2010. Lessons from the Global South: Reading and Learning with Technology, March 1, 2010.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8, 293–332.

Chasteen A. L., Park D. C., Schwarz N.(2001). Implementation intentions and facilitation of prospective memory. *Psychological Science*, 12, 457-461.

Deacon, H.S, P. N. Bowers, and J.R. Kirby. 2010. The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of Educational Research*, 80, 2. 144-179.

Dowrick, P. W. (2012). Self model theory: Learning from the future. *WIREs Cognitive Science.*

Dowrick, P. W., Weol Soon. Kim-Rupnow, and Thomas J. Power. 2006. Video Feedforward for Reading. *Journal of Special Education*, 39, 194-207.

Dowrick, P. W., Power, T. J., Ginsburg-Block, M., Kim-Rupnow, W. S., & Manz, P. H. (2000). ACE Reading: Feedforward to promote reading, communication, and problem-solving for students with disabilities (Final Report #H180G70001). Washington, DC: U.S. Department of Education.

Dowrick, P. W. 1999. A review of self modeling and related interventions. *Applied and Preventive Psychology* 8:23-39.

Dowrick P.W. Kim-Rupnow W.S. & Power T.L.2006. *Video feedforward for reading,* 39 Journal of Special Education 194-207.

Du Plessis, Joy and Irfan Muzaffar. 2010. Professional Learning Communities in the Teachers' College: A resource for teacher educators. USAID (EQUIP1), American Institutes of Research.

Feldon, David F. 2007. Cognitive Load and Classroom Teaching: The Double-Edged Sword of Automaticity. *Educational Psychologist*, 42(3), 123–137

Fisk, A. D. 1989. Training consistent components of tasks: Developing an instructional system based on automatic/controlled processing principles. Human Factors, 31, 453–463.

Fogassi, L., & Ferrari, P. F. (2010). Mirror systems. WIREs Cognitive Science, 2, 22-38. DOI: 10.1002/wcs.89

Frey, Scott and Valerie E. Gerry. 2006. Modulation of Neural Activity during Observational Learning of Actions and Their Sequential Orders. The Journal of Neuroscience, December 20, 2006, 26(51):13194-13201 (Watching With Intent To Repeat Ignites Key Learning Area Of Brain. ScienceDaily. http://www.sciencedaily.com-releases/2006/12/061222093112.htm

Gagné, M. & Deci, E. L. 2005. Self-determination Theory and Work Motivation. *Journal of Organizational Behavior*, 26, 331-362.

Gilbert, D. T., & Wilson, T. D. 2007. Prospection: Experiencing the future. Science, 317, 1351-1354.

Gollwitzer,P. M. and P. Sheeran. 2006. Implementation intentions and effective goal pursuit. A meta-analysis of effects and processes. *Advances in Experimental Social Psychology*, 38, 69-119.

Graetz, J. E., Mastropieri, M.A., & Scruggs, T.E. (2005). Show time: Using video self-modeling to decrease inappropriate behaviors. Teaching Exceptional Children, 38, 43-48. Gresham, F.M. (1998). Social skills training: Should we raze, remodel, or rebuild? *Behavioral Disorders*, 24, 19-25.

Gresham, F.M., Sugai, G, & Horner, R.H. (2001). Interpreting outcomes of social skills training for students with high incidence disabilities. *Exceptional Children*, 67, 331-344.

Hitchcock, C. H., Dowrick, P. W., & Prater, M. A. (2003). Video self-modeling intervention in school-based settings: A review. *Remedial and Special Education*, 24, 36-45.

Hitchcock, Caryl H., Mary Anne Prater, and Peter W Dowrick. 2004. Reading Comprehension and Fluency: Examining the Effects of Tutoring and Video Self-Modeling on First-Grade Students with Reading Difficulties. *Learning Disability Quarterly* 27 (2), 89-103.

Ho, Jennifer and Hetal Thukral. 2009. Tuned in to Student Success: Assessing the Impact of Interactive Radio Instruction for the Hardest-to-Reach. Washington, DC: Education Development Center.

IEG. 2007. Using Training to Build Capacity for Development: An Evaluation of the Project-Based and WBI Training. World Bank.

Jeannerod, M. 1994. The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, 17 (2), 187-202

Joyce, Bruce & Showers, Beverly. Student Achievement through Staff Development, 3rd Ed., ASCD, 2002.

Kliegel M., M. Martin, M.A McDaniel, G. O. Einstein, and C. Moor. 2007. Realizing complex delayed intentions in young and old adults: The role of planning aids. *Memory & Cognition*, 35, 1735-1746.

Knäuper B, A.Collam, A. Rosen-Brown, J. Lacaille, E. Kelso, M. Roseman. 2011. Fruitful plans: Adding targeted mental imagery to implementation intentions increases fruit consumption. *Psychology and Health*. Feb, 18:1-17

Liu, L.L., and D.C. Park. 2004. Aging and Medical Adherence: The Use of Automatic Processes to Achieve Effortful Things. *Psychology and Aging,* 19: 318–325.

Llambiri, S. 2005. "Braktisja e fshehtë", "Nga fundi në fillim, Tirana, Shtëpia Botuese "Albas"

Manderlink, G. and J. Harackiewicz. 1984. Proximal versus distal goal setting and intrinsic motivation. *Journal of Personality and Social Psychology*, 47 (4) 918.

McDaniel, Mark A. and Michael K. Scullin. 2010. (Implementation intention encoding does not automatize prospective memory responding. *Memory & Cognition*, 38, 221-232.)

Meharg, S. and Woltersdork, M.A. 1990. Therapeutic use of videotape self-modeling: A review. *Advances in Behavior Research and Therapy,* 12, 85-99.

Ondobaka, S., F. P. de Lange, R. D. Newman-Norlund, M. Wiemers, H. Bekkering. Interplay Between Action and Movement Intentions During Social Interaction. *Psychological Science*, 2011; 23 (1): 30 DOI: 10.1177/0956797611424163

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169-192.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2006). Mirrors in the mind. *Scientific American*, November, pp. 54-61.

Powel, Douglas R., Karen E. Diamond, Margaret R. Burchinal, and Matthew J. Koehler. 2010. Effects of an Early Literacy Professional Development Intervention on Head Start Teachers and Children. *Journal of Educational Psychology*, 102, 299–312.

Rizzolatti, G., L. Fadiga, L. Fogassi, and V. Gallese. 2002. "From Mirror Neurons to Imitation: Facts and Speculations." In The Imitative Mind: Development, Evolution, and Brain Bases, A.N. Meltzoff and W. Prinz, eds. Cambridge: Cambridge University Press.

Schultheiss, O. C., & Brunstein, J. C.1999. Goal imagery: Bridging the gap between implicit motives and explicit goals. *Journal of Personality*, 67, 1-38.

Smith, S. 1979. Remembering in and out of context. Journal. Of Experimental Psychology: Human *Learning and Memory*, 5, 460-471.

Smith, S. M., A. Glengerg, and R. A. Bjork 1978. Environmental context and human memory. *Memory and Cognition*, 6, 342-353.

Sweller, J. 1999. Instruction design in technical areas. Camberwell: Australian Council for Educational Research.

Ten Eyck, L. L., Gresky, D. P., & Lord, C. G. (2007). Effects of Directed Thinking on Exercise and Cardiovascular Fitness. *Journal of Applied Biobehavioral Research*, 2007; 12 (3-4): 237 DOI: 10.1111/j.1751-9861.2008.00023.x

UNESCO. 1998. World Education Report 1998: Teachers and Teaching in a Changing World. Paris: UNESCO.

Van Gog, Tamara, Liesbeth Kester, and Fred Paas. 2010. Effects of concurrent monitoring on cognitive load and performance as a function of task complexity. *Applied Cognitive Psychology*, June 24. http://dx.doi.org/10.1002/acp.1726

van Osch, Liesbeth, Lilian Lechner, Astrid Reubsaet, Stefan Wigger, Hein de Vries. 2008. *British Journal of Health Psychology*, 13 (3), 525-535.

Wendlinsky, H. (2002). The link between classroom teaching practices and student academic performances. *Education Policy Analysis* 10 (12) 1-64.

Whitsett, D.A., Forrest Dolgener, and Tanjala Mabon Kole. 1998. The Non-Runner's Marathon Trainer. New York: McGraw-Hill.

Wright, C. & Smith, D. (2009). The effect of PETTLEP imagery on strength performance. *International Journal of Sports and Exercise Psychology, 7*, 18-31.

Zhou, Jiangyuan. (2012) Effects of reciprocal imitation on teacher-student relationships and student learning outcomes. *Mind, Brain, and Education*, 6 (2), 66-73.

Ziegelmann, Jochen P., Sonia Lippkea and Ralf Schwarzera 2006.   Adoption and maintenance of physical activity: Planning interventions in young, middle-aged, and older adults Psychology & Health, 21(2), 145 – 163.

Ziegelmann, J.P., A. Luszczynska, S. Lippke, & R. Schwarzer. 2007. Are goal intentions or implementation intentions better predictors of health behavior? A longitudinal study in orthopedic rehabilitation. *Rehabilitation Psychology*, 52(1), 97–102.