Behavioral Development Bulletin
2016, Vol. 21, No. 1, 44–49

1942-0722/16/$12.00 http://dx.doi.org/10.1037/bdb0000018

Reinforcement of Correct Answers Raised Stage of Performance in Traditional Nonliterate Nepalese Adults

Dristi Adhikari
Dare Institute, Cambridge, Massachusetts

Thirty-three nonliterate Nepalese adults responded to the 2 stage-based "isolation of variables" (Inhelder & Piaget, 1958) instruments: the thatched roof problem and the laundry problem. The thatched roof instrument was used as the training instrument and administered individually. The laundry instrument was used as transfer task instrument and administered in a group. With the laundry instrument, all correct answers were reinforced with points and a monetary reward of 0.2 cents for Primary and Concrete stages; and 0.3 cents for Abstract and higher stages. The mean stage of performance increased from *M* stage 8.85 (*SD* = 0.86) to *M* stage 10.00 (*SD* = 1.07) from training instrument to the end of the transfer task training and testing, $t(32) = 7.60$, $p = .000$. This is roughly 1 stage increase from pretest to posttest. The frequency of people performing at the Primary Stage 8 and Concrete Stage 9 decreased at posttest. The frequency of people performing at Abstract Stage 10 and Formal Stage 11 increased at posttest. This shows that training with reinforcement has most productive effect on increasing stage performance from pretest to posttest. This outcome strongly suggests that all testing should include repeated presentation of very alike items and that reinforcement needs to be provided for all correct answers.

Keywords: cross-culture, feedback, pretest and posttest, reinforcement, stage change

Psychological assessment of "capability" across cultures is a difficult process. It is often muddled by ethnocentric content-laden approaches used to assess cognitive abilities of people in different cultures. This has been a steady issue in evaluating cognitive capacity of people in non-Western underdeveloped countries (Upadhyaya, Giri, & Commons, 2015). The results of such approaches in understanding cross-cultural capability underscores the actual possible performance. There is a necessity to make assessment approaches more applicable by drawing from real-life experiences of target populations while designing studies in psychology. Until now, it was assumed that changing content would be enough in making a study culturally relevant. The study design in the current paper is content-centric and consistent with previous studies, such as moral dilemma problems in Mexicali (Commons, Galaz-Fontes, & Morse, 2006), tool problems in Nepal (Giri, Commons, & Tuladhar, 2014) and reinforced correct answers in traditional non-literates found in the world (Upadhyaya et al., 2015).

The previous studies conducted with laundry and thatched roof have led to our understanding that content-centric study alone is not sufficient to precisely assess cross cultural population (Upadhyaya et al., 2015). Commons and Davidson (2015) presented that reinforcement of accurate responses on frequent task trials have significant effect on the achievement of performance transfer skills between nonrelevant and unfamiliar tasks. The conclusion of the current study also strongly advises that factors independent of the content materials, such as reinforcement, need to be included in the study to safeguard effective performance. The current study is a follow-up to the Commons and Davidson (2015) and Upadhyaya et al. (2015) studies that

Michael Lamport Commons discussed with the author and provided notes on which this paper is partially based. The author also thanks Ashok Paudel, Ashraya Dawadi, Ram Sharma, and Subeksha Poudel for their help with data collection and participant recruitment.

Correspondence concerning this article should be addressed to Dristi Adhikari, Dare Institute, 234 Huron Avenue, Cambridge, MA 02138-1328. E-mail: dristiadhikary@gmail.com

used cross culturally sensitive assessment instruments with reinforcement.

As in Commons and Davidson (2015), this study shows that training helps in improving stage performance when correct responses are reinforced. All correct responses in the transfer task were reinforced with money and points. In the transfer task, although each participant answered the questions individually like in thatched roof, all correct answers were reinforced with money and points. The points they received for correct answers were added to a group total. Johnson, Maruyama, Johnson, Nelson, and Skon (1981) found that this process of having participants work at discrete task toward a group reinforcement was significantly effective in successful individual performance.

It has been shown that all training helps in raising the measured stage of performance in a new domain or with new task content. In addition, it is known that actual stage change in Piagetian terms has been shown to be a function of $\log_2$age, and is hence a very slow process (Commons, Miller, & Giri, 2014). Therefore, when stage change occurs within a very short of time (<1 month), it may be because the participants may already have a higher stage of performance in some other domain or task content which when trained, decodes to the task at hand (Upadhyaya et al., 2015). We predicted quick stage change with the reinforcement of all correct answers (Commons & Davidson, 2015).

Method

Participants

Thirty-three nonliterate residents from two rural villages in eastern Nepal participated in the study. The illiteracy rate in the eastern part of Nepal is 40.40% (Bajracharya & Khaka, 2014). Their age ranged from 17 to 78 years, $M = 38.36$, $SD = 17.16$. The participants had never been to school. They never received any formal or informal education or training. They did not know how to read and write.

Instruments

Two stage based isolation of variables instruments were used in the study. The initial training instrument was called the Thatched Roof Instrument. This instrument was based on the Laundry problem (Commons et al., 2014). The laundry instrument is derived from Inhelder and Piaget's (1958) pendulum problem. Participants were asked to detect causal associations from various systems and then compare the systems. These simple causality detecting problem instruments were put into behavioral developmental form. They extended down to early Primary Stage 8 and up to Metasystematic Stage 13.

For the training instrument, the task was based on building a thatched roof in eight different ways. The participants were told the roof would either be strong or weak based on the multiple variables: concrete roof, slate roof; yellow hay, green hay; tin sheet, plastic sheet; thick twine with wooden frame, thick twine. Participants were asked to identify a single causal variable from the different possible pairs of variables that formed the result of either strong or weak roof. Example of the thatched roof problems is shown in Table 1.

The thatched roof was used as the training instrument. Administered individually, the thatched roof instrument was operated to help the participants achieve their highest stage of performance from other spheres to the task of detecting casual variable. The participants were not reinforced with money or points. Each stage was made up of a number of tasks. The stages ranged from Primary Stage 8 to Metasystematic Stage 13. The participants responded to the instrument in the ascending order of stage. Primary Stage 8 tasks were asked first. The decision rule of operating thatched roof was identical to Upadhyaya et al. (2015). To advance to a higher stage, the participants had to answer three questions correctly in a row without a time limit. So if a participant answered three questions correctly in a row for Primary Stage 8, they would advance to Concrete Stage 9. However, if a participant answered only two questions correctly in a row, they would be presented from task one stage lower. If a participant got two answers correct but the third incorrect, they would not move forward until all three were correct. For example,

> if a participant did not get three consecutive corrects for Formal Stage 11, and then the task would be moved down a stage to Abstract Stage 10. They would have to answer three consecutive Abstract Stage 10 tasks correctly before moving to Formal Stage 11 again. (Upadhyaya et al., 2015, p. 71)

Table 1
Example of the Thatched Roof Problem

There are six ways a roof can be built. Sometimes it will be strong after being built and sometimes it will be weak.					
Concrete roof	Thick twine	Tin sheet	Yellow hay	→	Weak
Slate roof	Thick twine + Wooden frame	Plastic Sheet	Green hay	→	Strong
Concrete roof	Thick twine	Plastic Sheet	Green hay	→	Weak
Slate roof	Thick twine	Plastic Sheet	Yellow hay	→	Weak
Concrete roof	Thick twine + Wooden frame	Tin Sheet	Green hay	→	Strong
Slate roof	Thick twine + Wooden frame	Tin Sheet	Yellow hay	→	Strong
Look back at the examples. After being built, will the roof be strong or weak?					
Slated roof	Thick twine	Tin sheet	Green hay	Strong	Weak
Concrete roof	Thick twine + Wooden frame	Tin sheet	Yellow hay	Strong	Weak
Concrete roof	Thick twine	Plastic sheet	Yellow hay	Strong	Weak
Slated roof	Thick twine + Wooden frame	Tin sheet	Green hay	Strong	Weak
Slated roof	Thick twine	Tin sheet	Yellow hay	Strong	Weak
Slated roof	Thick twine	Plastic sheet	Green hay	Strong	Weak
Concrete roof	Thick twine + Wooden frame	Plastic sheet	Green hay	Strong	Weak
Concrete roof	Thick twine	Tin sheet	Green hay	Strong	Weak
Slated roof	Thick twine + Wooden frame	Plastic sheet	Yellow hay	Strong	Weak
Concrete roof	Thick twine + Wooden frame	Plastic sheet	Yellow hay	Strong	Weak

At higher stages, Systematic Stage 12 and Metasytematic Stage 13, it was expected that the participants may not get three consecutive correct answers in a row. Hence, feedback was given on how to identify the casual variables.

After operating the thatched roof, the participants were divided into six groups and given the transfer task instrument, the laundry instrument. An example of the laundry problem is shown in Table 2 and 3.

Procedure

The laundry instrument was operated after the completion of the thatched roof problem. Identical to the thatched roof problem, the laundry transfer task instrument had tasks ranging from the Primary Stage 8 to Metasystematic Stage 13. Participants were given the problems in the ascending order of stage.

The decision rule of administering the laundry instrument was similar to thatched roof. However, laundry instrument was administered in a group. The correct answers were reinforced with money and points. The points that individuals received for correct answers were added to a group total. The group with most cumulative points won an additional bonus amount of Nepalese rupees 600 ($6).

For Primary Stage 8 and Concrete Stage 9, the participants would have to give three consecutive correct answers to advance to higher stage. If the participants did not give three consecutive correct answers, they were asked the two tasks again and reinforced for each correct

Table 2
Example of the Laundry Problem (1977)

A cloth was stained with red lipstick. There are six ways it can be washed. Sometimes it will be clean after being washed and sometimes it will be dirty.					
Brand A bleach	Powdered soap	Blue booster	Cold water	→	Dirty
Brand B bleach	Liquid soap	Pink booster	Hot water	→	Clean
Brand A bleach	Powdered soap	Pink booster	Hot water	→	Dirty
Brand B bleach	Powdered soap	Pink booster	Cold water	→	Dirty
Brand A bleach	Liquid soap	Blue booster	Hot water	→	Clean
Brand B bleach	Liquid soap	Blue booster	Cold water	→	Clean

Table 3
Six Sample Test Tasks of the Possible Ten of the Laundry Problem

Look back at the examples. After being washed, will the cloth be clean or dirty?					
Brand B bleach	Powdered soap	Blue booster	Hot water	→	Clean or dirty?
Brand A bleach	Liquid soap	Pink booster	Cold water	→	Clean or dirty?
Brand A bleach	Powdered soap	Blue booster	Cold water	→	Clean or dirty?
Brand B bleach	Powdered soap	Pink booster	Hot water	→	Clean or dirty?
Brand B bleach	Liquid soap	Pink booster	Hot water	→	Clean or dirty?
Brand A bleach	Liquid soap	Blue booster	Cold water	→	Clean or dirty?

response. Every reinforcement was a monetary reward of 2 Nepalese rupees, which is about 0.2 cents (USD) and a point. If a participant did not get two Primary Stage 8 tasks correct in a row for two consecutive trials, then their stage was recorded as Preoperational Stage 7. The research team was asked to move on to testing a different participant if there were any participants in Preoperational Stage 7.

For participants to reach Formal Stage 11 and above, they would have to get four tasks correct in a row to move to the next stage tasks. Every reinforcer was a monetary reward of 3 Nepalese rupees which is about 0.3 cents (USD) and a point. If the participants did not get four tasks correct in a row after three consecutive trials, then the participant would move down in the stage of the task presented. For example, if a participant did not get four Abstract Stage 10 tasks correct in a row for three consecutive trials, then the participant would be moved one stage down to Concrete Stage 9. The participants would then be given primary stage Concrete Stage 9 tasks and would be reinforced monetarily for each correct answer. The participant would move back to Abstract Stage 10 tasks once the participant got three Concrete Stage tasks correct.

As in Upadhyaya et al. (2015), it was expected that participants may not get four consecutive correct answers at Systematic Stage 12 and Metasytematic Stage 13. In fact, they might not get the tasks correct when moved down one stage and monetarily reinforced. The highest stage was recorded as the stage at which participants got stuck and could not advance any further despite reinforcing each trial. The participants were shown how to approach high stage tasks after recording their highest stage (Upadhyaya et al., 2015, p. 73).

Results

Even though we analyzed the data using groups, this is essentially a single subject design in which each person served as their own control. The acquisition curves for the Laundry are available. Also if one looks at the initial stage at the beginning of training, that served as the control. The distribution of participant stage of performance on the pretest and posttest are shown (see Table 4 and Figure 1). This was done to test whether participant's performance enhanced at posttest. This was also done to see whether the occurrence of people performing at the lower stages decreased at posttest and the frequency of people performing at the higher stages increased at posttest.

The pretest stage was the thatched roof of the participants. This was the final stage they reached after moving down a stage for not answering three consecutive questions correctly. Posttest stage was the laundry instrument of the participants administered in a group after reinforcement of each correct answer. This was the last stage that participants reached after moving down a stage for not answering three consecutive questions correctly in lower stages and four consecutive questions correctly in higher stages.

The range of pretest and posttest stage performance differences are shown in Table 4, and the

Table 4
Pretest and Posttest Frequency for the Stages

Stage	Pretest frequency	Posttest frequency
8	14	2
9	11	9
10	7	12
11	1	8
12	0	1
13	0	1

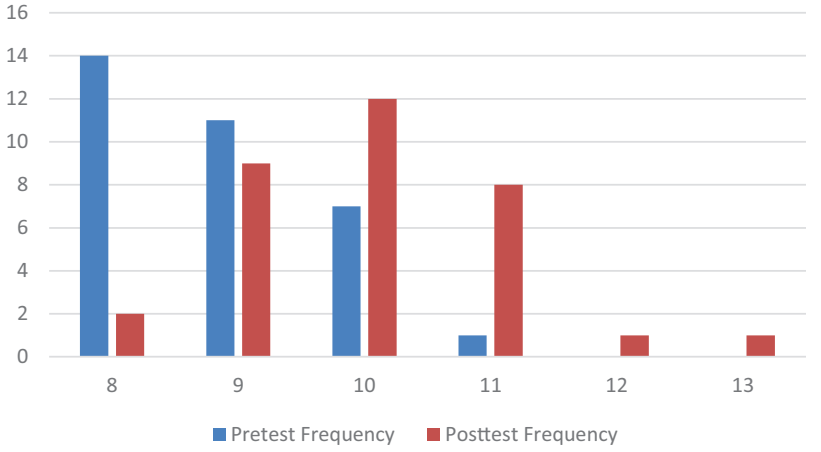

Figure 1. Pretest and posttest frequency for the stages.

frequency of the range of those differences is shown in Table 5. The table shows that 90.90% of the range data was concentrated in the increase of at least 1 stage. This again shows that the data do not have a lot of variation. Hence, there was also a strong correlation between the beginning stage and the stage someone ended up at.

A repeated-measure *t* test was run to see the differences between the pretest and posttest means. The mean stage increased from $M = 8.85$ ($SD = 0.86$) to $M = 10.00$ ($SD = 1.07$), $t(32) = 7.60$, $p = .000$, Cohen's $d = 1.18$. The size of the difference is roughly one stage from pretest to posttest. Whereas it is true that the overall mean increased, the frequencies also show that a number of people moved from lower stages to higher stages. There were 10 in total of individuals performing at Formal Stage 11, Systematic Stage 12, and Metasystematic Stage 13 in the posttest. There was only 1 individual performing at this stage in the pretest.

Table 5
Range of Pretest and Posttest With the Frequency of the Range

Valid	Frequency	Percent	Valid percent
0	3	9.1	9.1
1	20	60.6	60.6
2	9	27.3	27.3
3	1	3.0	3.0
Total	33	100	100

Even though individuals did move up in stage of performance, it was also true that individuals' posttest scores were predicted by their pretest score. Posttest scores were compared with pretest scores by running a simple linear regression. The pretest and posttest correlated with an $r = .613$, $p = 000$. Note that this is a significantly large *r*. This also shows that participants improved their performance by at least 1 stage.

Discussion

Reinforcement of correct answers produced significant stage change in performance among the nonliterates of eastern Nepal. Training, with reinforcement of correct answers, seems to have increased stage of the participants at least by one. Individuals who performed at the Concrete Stage 10 in the pretest were likely to perform at the Formal Stage 11. One might argue that this change could be the result of practice and not reinforcement. However, using just practice with feedback and no reinforcement had a negative effect on acquisition. With the laundry and thatched roof problem, when attempting the next stage beyond where one tests, one is initially wrong at first half the time. Lots of people just quit. That is about the same as receiving a D grade. That has been partially tested but in a much younger group (Commons & Davidson, 2015). Performance observed in this study was achieved because of the support provided (Fischer, Hand, & Russell, 1984) in the form of reinforced correct

answers and practice. This probably aided as an extra support for participants to carry out the task.

It is important to note that the participants improved their stage at least by one stage but only some advanced to higher stages (Systematic Stage 12 and Metasystematic Stage 13). Given the difficult conditions where there was only single trial per stage, it might be the case that people did not advance at all because they never acquired the lower stage skills necessary to progress. It might be that the amount of money offered for correct answers did not have enough value for them to adequately reinforce getting correct answers.

For the future studies, it is suggested that a more structured practice of sticking with the lower stage tasks for up to 4 trials be followed. It is also suggested that half the participants be run with laundry task and half with thatched roof as we did in this case. Through this we will be able to compare the acquisition slope as well as how long it takes before the acquisition starts to take place. Further, if the people do not improve by the third or fourth trial, larger amount of money might be tried after the rest of the participants have completed the study in that particular village.

As in a previous Nepal study (Upadhyaya et al., 2015), this finding powerfully recommends that all assessments should include frequent exhibition of very similar items and that correct answers should be reinforced. Otherwise there is a threat of undervaluing what tasks people can effectively accomplish and what their accurate stage of performance is.

References

Bajracharya, A., & Khaka, S. (2014). *Literacy status in Nepal*. Retrieved from http://unesdoc.unesco.org/images/0022/002276/227683e.pdf

Commons, M. L., & Davidson, M. N. (2015). The sufficiency of reinforcing problem solutions for transition to formal operations. *Behavior Development Bulletin, 20,* 114–130.

Commons, M. L., Galaz-Fontes, J. F., & Morse, S. J. (2006). Leadership, cross-cultural contact, socio-economic status and formal operational reasoning about moral dilemmas among Mexican nonliterate adults and high school students. *Journal of Moral Education, 35,* 247–267. http://dx.doi.org/10.1080/03057240600681785

Commons, M. L., Li, E. L., Richardson, A. M., Gane-McCalla, R., Barker, C. D., & Tuladhar, C. T. (2014). Does the model of hierarchical complexity produce significant gaps between orders and are the orders equally spaced? *Journal of Applied Measurement, 15,* 422–449.

Commons, M. L., Miller, L. S., & Giri, S. (2014). A model of stage change explains the average rate of stage of development and its relationship to the predicted average stage ("smarts"). *Behavioral Development Bulletin, 19,* 1–11. http://dx.doi.org/10.1037/h0101076

Fischer, K. W., Hand, H. H., & Russell, S. (1984). The development of abstractions in adolescence and adulthood. In M. Commons, F. A. Richards, & C. Armon (Eds.), *Beyond formal operations* (pp. 43–73). New York, NY: Praeger.

Giri, S., Commons, M. L., & Tuladhar, C. T. (2014, June). Stage of performance in nonliterate and unschooled Nepalese adults: A universal evolutionary-behavioral-developmental assessment based on knowledge of tool usage. Presented at the 29th Annual Symposium for the Society for Research in Adult Development, Salem, MA.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York, NY: Basic Books. http://dx.doi.org/10.1037/10034-000

Johnson, D. W., Maruyama, G., Johnson, R. W., Nelson, D., & Skon, L. (1981). Effects of cooperative, competitive, and individualistic goal structures on achievement: A meta-analysis. *Psychological Bulletin, 89,* 47–62. http://dx.doi.org/10.1037/0033-2909.89.1.47

Upadhyaya, R. E., Giri, S., & Commons, M. L. (2015). Reinforced correct answers to next stage problems produced the highest stage performance in traditional non-literates found in the world. *Behavioral Development Bulletin, 20,* 79–75.

Received March 2, 2016
Revision received March 18, 2016
Accepted March 18, 2016 ■