ASSESSING PROBLEM-SOLVING DISPOSITIONS: LIKELIHOOD-TO-ACT SURVEY

Kien Lim
University of Texas at El Paso
kienlim@utep.edu

Osvaldo Morera
University of Texas at El Paso
omorera@utep.edu

Mourat Tchoshanov
University of Texas at El Paso
mouratt@utep.edu

This paper reports an ongoing study that is aimed at developing an instrument for measuring two particular problem-solving dispositions: (a) impulsive disposition refers to students' proclivity to spontaneously proceed with an action that comes to mind, and (b) analytic disposition refers to the tendency to analyze the problem situation. The instrument is under development and consists of likelihood-to-act items in which participants indicate on a scale of 1 to 5 how likely they are to take a particular action in a given situation. The instrument was administered to 318 college students, mainly pre-service teachers. Statistical analysis indicates that likelihood-to-act items are reliable and that the current version of the instrument has room for further improvement.

Motivation for the Study

For many mathematics students, "doing mathematics means following rules laid down by the teacher, knowing mathematics means remembering and applying the correct rule when the teacher asks a question, and mathematical truth is determined when the answer is ratified by the teacher" (Lampert, 1990, p. 31). Students with such beliefs tend to exhibit dispositions such as "waiting to be told what to do," "doing whatever first comes to mind," and "diving into the first approach that comes to mind" (Watson & Mason, 2007, p. 207). In this paper, we use the term *impulsive disposition* to mean the tendency to "spontaneously proceed with an action that comes to mind without analyzing the problem situation and without considering the relevance of the anticipated action to the problem situation" (Lim, 2008a, p. 49).

Some problem-solving episodes found in mathematics education literature can be interpreted as instantiations of impulsive dispositions. For example, consider the following missing-value problem that was posed by Cramer, Post and Currier (1993) to pre-service teachers: *Sue and Julie were running equally fast around a track. Sue started first. When she had run 9 laps, Julie had run 3 laps. When Julie completed 15 laps, how many laps had Sue run?* Thirty-two out of 33 pre-service teachers solved this problem by setting up a proportion such as $9/3 = x/15$. These pre-service teachers are considered *impulsive* if they had applied the proportion algorithm without analyzing the problem situation. In fact, Lim (2008b) found that after a course on rational numbers and algebraic reasoning pre-service teachers, on average, performed better on all four direct-proportional problems but worse on all three non-direct-proportional problems.

As mathematics educators, we are interested in helping students advance from impulsive disposition to *analytic disposition*, in which a student "attempts to understand the problem statement, studies the constraints, identifies a goal, imagines what-if scenarios, and/or considers alternatives" (Lim, 2008a, p. 45). To track this advancement, we need to identify where a student stands in terms of his or her disposition. In other words, we need an efficient and reliable instrument that can "measure" students' impulsive disposition and analytic disposition. In the field of mathematics education, it appears that no such instrument has been developed. In this paper we present a few theoretical constructs related to impulsive disposition, overview the literature associated with assessing cognitive constructs through the use of survey, report our research process, and discuss the results that we have obtained.

Theoretical Constructs Related to Impulsive Disposition

Psychological perspective. In terms of cognitive tempo or response style, a person may be classified as either *impulsive* or *reflective*. Kagan, Rosman, Day, Albert, and Phillips (1964) constructed the Matching Familiar Figures Test to measure children's cognitive tempo. An impulsive is one whose response time is faster than the median and whose accuracy rate is below the median, whereas a reflective is one whose response time is slower than the median and whose accuracy rate is above the median. Nietfeld and Bosma (2003) describe impulsives as "individuals who act without much forethought, are spontaneous, and take more risks in everyday activities" (p. 119) whereas reflectives are "more cautious, intent upon correctness or accuracy, and take more time to ponder situations" (p. 119). In their study on consistency in cognitive responses among adults across academic tasks, Nietfeld and Bosma found moderate positive correlations for response styles among the three types of tasks they investigated: verbal, mathematical, and spatial. The mathematical tasks used in their study were two-digit addition or subtraction problems arranged in a traditional vertical format. Although such tasks are appropriate for measuring cognitive tempo along a speed-accuracy continuum, they are not appropriate for measuring disposition along an impulsive-analytic continuum. Whereas an impulsive tempo is characterized by a fast but inaccurate response, an impulsive disposition is characterized by "diving into the first approach that comes to mind" and not necessarily by how fast an approach comes to mind.

Problem-solving perspective. Schoenfeld (1985) has identified four categories of cognition that provide a framework for analyzing problem-solving behaviors: (a) mathematical knowledge base, (b) use of heuristics, (c) monitoring and control, and (d) beliefs about mathematics and doing mathematics. Impulsive disposition can be regarded as an externalization of certain beliefs such as "there is only one correct way to solve any mathematics problem—usually the rule the teacher has most recently demonstrated to the class" (Schoenfeld, 1992, p.359). Impulsive disposition can also be considered as a lack of metacognition—a term introduced by Flavell (1976) as "the active monitoring and consequent regulation and orchestration of these processes in relation to the cognitive objects or data on which they bear, usually in the service of some concrete goal or objective" (p. 232).

Teaching-learning perspective. According to Harel (2008), mathematics consists of two complementary sets: (a) *ways of understanding* refer to the products of mental acts while doing mathematics; they include definitions, theorems, proofs, problems, and solutions, and (b) *ways of thinking* refer to the characteristics of the mental acts while doing mathematics. Harel (2007) stipulates that "students develop ways of thinking only through the construction of ways of understanding, and the ways of understanding they produce are determined by the ways of thinking they possess" (p. 272). According to this principle, it is counter-productive to help students develop ways of understanding without helping them develop ways of thinking, and vice versa. Hence, students should be provided opportunities to engage in mental acts (e.g., generalizing, justifying, problem-solving, symbolizing, computing, generalizing, predicting, etc.) that can advance both their ways of understanding and ways of thinking. Lim (2008a) identifies *impulsive anticipation* and *analytic anticipation* as two ways of thinking in the context of problem solving. An important goal of mathematics education is to help students advance from undesirable ways of thinking (e.g., impulsive disposition, authoritative proof scheme) to desirable ways of thinking (e.g., analytic disposition, deductive proof scheme).

Means for Assessing Impulsive Disposition and Analytic Disposition

One useful way to measure cognitive and psychological constructs is through the use of survey development. The use of surveys can be very informative, as they allow for the quantification of the constructs under study. With such quantification, we can investigate group differences on those constructs and assess how those constructs associate with other behavioral measures. Examples of the use of such measures in the psychological literature are vast, ranging from the measurement of social problem solving (D'Zurilla, Nezu, & Maydeu-Olivares, 2002) to the measurement of decision making styles (Nygren, 2000).

Nygren (2000) constructed the Decision Making Styles Inventory, which measures the degree to which a person makes everyday decisions using an analytical approach, an intuitive approach and an approach which minimizes regret. Analytical decision making involves considering every aspect of the problem before making a decision whereas intuitive decision making involves a reliance on one's gut feeling. These two constructs are analogous to analytic disposition and impulsive disposition.

The goal of this project was to develop a measure of mathematical disposition. We wanted to demonstrate the internal consistency reliability of the survey items. In addition, we wanted to see how well the items in each subscale are inter-correlated, and how well the two items in each impulsive-analytic pair are correlated. Finally, we wanted to assess how scores on such a measure are related to self-reported academic performance in mathematical classes and the participant's teacher training program.

Research Process

Instrument design, testing, and refining are an elaborate process involving multiple cycles.

Survey Development

The initial instrument designed to assess impulsive disposition was a multiple-choice test on ratios and proportions. Students have a tendency to overuse proportional strategies for solving missing-value problems (Van Dooren, De Bock, Hessels, Janssens, & Vershaffel, 2005). The items were designed to determine whether students inappropriately use a proportion to solve missing-value problems that do not involve a direct-proportional situation (e.g., an additive situation, an inverse-proportional situation) or inappropriately use a ratio to compare "non-rate" quantities (e.g., the size of a person's palm, the magnitude of a project in terms of worker-hours).

In subsequent versions multiple-choice items were used for students to choose the action that they would most likely perform in a given scenario. The format was eventually changed from a multiple-choice test to a likelihood-to-act survey which takes less time for students to complete.

The likelihood-to-act survey developed for this ongoing study has undergone two revisions. The first two exploratory versions were administered to about 70 pre-service middle-school teachers and 14 graduate students (mainly in-service teachers) respectively in courses taught by the first author. The version reported in this paper consists of nine pairs of *likelihood-to-act* items, sequenced from A to R. Four pairs (A-J, N-E, O-F, and I-R) involve equation solving; two pairs (B-K and L-C) involve word problems; two pairs (D-M and Q-H) involve fraction division and fraction addition respectively; and one pair (G-P) involves geometry. Figure 1 shows 3 pair of such items.

Two versions of the likelihood-to-act survey were used in this study. All the items in Version

1 are considered "specific" items in that a specific scenario is provided. The nine impulsive-disposition items are based on specific mathematical rules, formulas, or procedures that are supposedly familiar to students.

Please indicate, as honestly as you can, how likely you are going to act in the manner specified in the statement using the following scale:

1. Extremely Unlikely 2. Unlikely 3. Not Sure 4. Likely 5. Extremely Likely

B. When asked to find the cost of 18 cans of specialty soda given that 6 cans of specialty soda cost $4.10, how likely are you going to begin by setting up a proportion?

K. When asked to find the cost of 20 bottles of mineral water given that 4 bottles cost $2.10, how likely are you going to study the values of the quantities and predict the answer?

D. When you are asked to find the answer for $\frac{55}{95} \div \frac{11}{95}$ without using a calculator, how likely are you going to use the invert-and-multiply rule?

M. When you are asked to find the answer for $\frac{44}{82} \div \frac{11}{82}$ without using a calculator, how likely are you going to study the two fractions and predict the answer?

I. When asked to simplify an equation (e.g., $10^{3x} \cdot 10^{2y} = 1000 \cdot 10^{3x}$), how likely are you going to begin by applying a formula (e.g., $B^m \cdot B^n = B^{m+n}$) or by following a procedure (e.g., taking $\log_{10}$ on both sides)?

R. When asked to solve $2^{5a} \cdot 2^{10b} = 8 \cdot 2^{5a}$ for b, how likely are you going to begin by inspecting the terms in the equation?

Figure 1. Three pairs of "specific" likelihood-to-act problems in Version 1.

Version 2 differs from Version 1 in that it contains 2 pairs of "general" items (see Figure 2); the remaining 7 pairs are identical to those in Version 1. "General" items were found to be less reliable in the pilot testing of an earlier version based on a small sample of 14 students. Version 2 was developed to verify this finding. Note that Pair I'-R' is analogous to Pair I-R, but Pair G'-P' is substantially different from Pair G-P.

G'. When asked to solve a word problem, how likely are you going to begin by applying a formula or using a procedure that comes to mind?

P'. When asked to solve a word problem, how likely are you going to begin by identifying the quantities in the problem and thinking about how the quantities are related?

I'. When asked to solve an equation, how likely are you going to begin by using a procedure (e.g. combining like terms) or searching for a formula (e.g. $B^m \cdot B^n = B^{m+n}$)?

R'. When asked to solve an equation, how likely are you going to inspect the terms in the equation before applying a standard procedure for solving the equation?

Figure 2. Two pairs of general likelihood-to-act problems in Version 2.

Data Collection and Analysis

A survey was administered in 13 mathematics classes in the final week of classes of the Fall 2008 semester. To encourage participation, a participant in each class was randomly selected to win a $10 gift voucher. The survey is comprised of two parts: (a) 18 likelihood-to-act items, and (b) either 18 *need-for-cognition* items or 18 *belief-attitude-confidence-in-algebra* items (these items are not discussed in this paper). 318 students were administered the survey, with 257 participants from 10 classes taking Version 1 and 61 participants from 3 classes taking Version 2 of the likelihood-to-act part.

Inter-item correlations were computed using Pearson correlations for the nine impulsive-disposition items and the nine analytic-disposition items. Items that are not significantly correlated with other items in the same category were analyzed to see if they could be improved or should be excluded from the next version of the instrument. Because the items were paired, the correlation between the impulsive-disposition item and analytic-disposition item in each pair was also determined. The reliability for each sub-scale of seven common items was determined using Cronbach's Alpha coefficient based on all 318 individuals. To assess the validity of each subscale, we also performed a 4×3 analysis of variance—four programs (Early Childhood to Grade 4 Generalist program, Grades 4-8 Generalist program, Grades 4-8 Math Specialist program, and B.S. Math program) by three self-reported grade-point-averages for mathematics courses (A, B, and C or below).

Results and Discussion

Reliability of the likelihood to act measure. To compute the reliability of the likelihood-to-act subscales, we used the seven pairs of items common to both versions to create a larger sample—318 individuals. Missing data on any of these items for these 318 students were imputed. At most, one item had seven missing item responses. The Cronbach's Alpha estimate of internal consistency reliability for the seven impulsive-disposition items was 0.64, (95% Confidence Interval: 0.58, 0.70). The reliability of the seven analytical items was 0.63 (95% Confidence Interval: 0.56, 0.69). While these reliability estimates are not very high, it should be noted that each subscale has only seven items. It is well known from classical test theory that the addition of items tends to increase test score reliability. For the next phase, we will focus on developing and testing additional items, s well as improving existing items.

Impulsive-disposition items. The correlations among the nine general impulsive-disposition items in Version 1 are presented in Table 1 (ignore the two specific items, G' and I', for the time being). By excluding items L, N, and I, all the correlations among the six remaining items (A, B, D, O, Q, and G) are significant with $p < 0.01$. Items L, N, and I will be replaced in the next version of the instrument. The strong correlations among items A, B, D, O, Q, and G suggest that impulsive disposition is a trait that cuts across the four domains: equation-solving (A and O), word problem (D and Q), fractions (B), and geometry (G).

Table 1
Correlations among the Nine Impulsive-Disposition Items

Item	A	B	L	D	N	O	Q	G	I	G’	I’
A	1										
B	0.32**	1									
L	0.15**	0.15**	1								
D	0.37**	0.26**	0.18**	1							
N	0.13*	0.17**	0.21**	0.10	1						
O	0.30**	0.30**	0.14*	0.29**	0.22**	1					
Q	0.25**	0.17**	0.09	0.15**	0.12*	0.21**	1				
G	0.24**	0.31**	0.13*	0.19**	0.17**	0.28**	0.23**	1			
I	0.02	0.15*	0.09	0.14*	0.26**	0.22**	-0.04	0.18**	1		
G’	0.34**	0.18	0.42**	0.07	0.36**	0.02	0.25	-	-	1	
I’	0.47**	0.42**	0.28*	0.29*	0.09	0.49**	0.33**	-	-	0.19	1

Note. The *p*-values for the correlation of 0.29 for I’ and D is lower than that for the correlation of 0.29 for O and D because the sample size was 61 for I’ and D (Version 2) and 315 for O and D (Versions 1 and 2).
$*p < .05$. $**p < .01$.

As for Version 2 (G’ and I’ instead of G and I), we should retain items A, B, D, O, Q, and I’ and replace items L, N, and G’. Interestingly, the general item I’ appears to be better correlations than the specific item I. A probable explanation is that the participants might have difficulty interpreting Item I because of the equation $10^{3x}\cdot10^{2y} = 1000\cdot10^{3x}$ and the meaning of $\log_{10}$.

Analytic-disposition items. The correlations among the 11 analytic-disposition items shown in Table 2 are generally less significant when compared to those in Table 1. This finding suggests that the analytic-disposition items are not as effective as the impulsive-disposition items. Items C, E, P, R, P’ and R’ have to be excluded in order for the remaining correlations to be significant. However, items J, K, M, F and H can remain intact for the next version. Hence, the individual items in these five pairs, A-J, B-K, D-M, O-F, and Q-H, seem to be reliable.

Table 2
Correlations among the 9 Analytic-Disposition Items

Item	J	K	C	M	E	F	H	P	R	P’	R’
J	1										
K	0.28**	1									
C	0.05	0.14*	1								
M	0.31**	0.41**	0.07	1							
E	0.17**	0.05	0.08	0.08	1						
F	0.22**	0.16**	0.21**	0.17**	0.19**	1					
H	0.26**	0.39**	0.00	0.43**	0.07	0.23**	1				
P	0.02	0.09	0.14*	0.19**	0.12*	0.10	0.06	1			
R	0.21**	0.18**	0.10	0.14*	0.23**	0.08	0.08	0.34**	1		
P’	-0.01	-0.14	0.21	-0.05	0.35**	0.06	-0.33**	-	-	1	
R’	0.20	0.07	0.44**	0.09	0.15	0.31*	-0.01	-	-	0.35**	1

Note. **p* < .05, ***p* < .01.

Correlation between the two items in each pair. The last column in Table 3 shows the correlation between the impulsive-disposition item and the analytic-disposition item in each pair. A significant negative correlation in Pair Q-H indicates that this pair of items differentiates impulsive disposition from analytic disposition. The significant positive correlations in Pair L-C and in Pair I-R mean that these two pairs of items should not be used in the next version. Note that the five good pairs (A-J, B-K, D-M, O-F, and Q-H) have either negative correlations or very small positive correlations. The lack of significant negative correlations suggests the possibility that impulsive disposition and analytic disposition are not necessarily mutually exclusive. In other words, a person may have two competing dispositions for the same problem situation.

Table 3
Comparing the Two Items in Each Pair

	Mean for the impulsive item	Mean for the analytic item	Difference betw. the two means	Correlation betw. the two items
Pair A-J	3.62	3.53	0.10	0.02
Pair B-K	3.66	3.53	0.12	-0.10
Pair L-C	3.34	3.55	-0.21	0.19*
Pair D-M	4.19	3.21	0.99	-0.00
Pair N-E	3.14	3.55	-0.40	0.10
Pair O-F	3.77	3.53	0.24	0.04
Pair Q-H	4.00	3.33	0.67	-0.16**
Pair G-P	3.40	3.84	-0.44	0.12
Pair I-R	3.77	4.09	-0.32	0.14*
Pair G’-P’	3.86	4.07	-0.21	0.24
Pair I’-R’	3.37	4.28	-0.91	0.16

Note. **p* < .05, ***p* < .01.

Association between training program, self-reported mathematics grade, and likelihood-to-act scores. To assess the validity of the likelihood-to-act measure, we also performed a 4 (programs) by 3 (self-reported numerical grade) analysis of variance on the analytical and impulsive composite scores. For the analytical subscale, there was a main effect for program ($F(3,270) = 8.233$, $p < 0.001$, $\eta 2 = 0.08$). Interestingly, elementary-school generalists (EC-4 program) had higher analytical scores than middle-school math specialists (4-8 Math program) and mathematics majors (B.S. Math program). Middle-school generalists (4-8 Generalists program) also had higher analytical scores than mathematics major. An interaction between self-reported grade and training program emerged only for the 4-8 Math program. Surprisingly, individuals in the 4-8 Math training program who reported a letter grade of B had higher analytical scores than those who self-reported a letter grade of A in their math coursework.

A similar analysis was performed for scores on the impulsive-disposition measure. There was a main effect for training program ($F(3,270) = 4.872$, $p = 0.003$, $\eta 2 = 0.05$). Math majors had higher impulsive-disposition scores than did all other groups. There were no differences among the other groups. There was no main effect for self-reported grade and no interaction between self-reported grade and training program. In summary, these results seem to suggest that students with increased exposure to traditional math coursework are less analytical and more impulsive.

Conclusion

The findings obtained in this study confirm the viability of using likelihood-to-act items to measure impulsive disposition and analytic disposition. Five out of eleven pairs of items have high inter-correlations and will be retained in the next version. The weaker items will be refined for the next version of the instrument. The reliability for the two subscales are 0.64 and 0.63. Our goal is to continue improving the instrument until a reliability of at least 0.75 is obtained.

Analytic-disposition items were found to be slightly less reliable than impulsive-disposition items due in large part to the former being typically less clear than the latter. The specific procedure or rule to which students are drawn can be stated explicitly in an impulsive-disposition item, but not in an analytic-disposition item. There is insufficient evidence in this study to support the claim that general items are not as reliable as specific item. Results from the 4-by-3 analysis of variance reveal an unexpected phenomenon. Students in more advanced but traditional math programs (i.e., B.S. program) were found to have lower analytic-disposition scores than those in the less advanced but reform-oriented mathematics programs (i.e., EC-4 program and 4-8 Generalists program). Future research is needed to account for this phenomenon.

One of the advantages of the likelihood-to-act survey is that it takes less time for students to complete than a mathematics test. Whereas students need to solve a problem in order to arrive at an answer choice in a test item, students only need to understand the problem statement and the action for consideration in a likelihood-to-act item to choose from a scale of 1 to 5 the likelihood level. Another advantage is that participants are less likely to feel threatened because the instrument, as a survey, is not perceived as an assessment of their mathematical knowledge.

However, like any survey, what participants say they will do may differ from what they actually do in a mathematics assessment or in a problem-solving situation. This raises the issue of the validity of the likelihood-to-act survey. Another limitation of the instrument is that an impulsive item is effective only if the students are familiar with the particular rule, formula or procedure that is mentioned in the item. For example, Item B will not be valid if it is administered to an elementary student who has not learned how to set up a proportion. Hence, the validity of a likelihood-to-act survey is limited to the group of students for which it is designed.

The likelihood-to-act items developed in this study were aimed at measuring impulsive disposition and analytic disposition. The idea of asking participants to indicate their likelihood to act may be extended to measure other dispositions such as waiting to be told what to do, relying on the teacher, consulting with peers, and so forth.

References

Cramer, K., Post, T., & Currier, S. (1993). Learning and teaching ratio and proportion: Research implications. In D. T. Owens (Ed.), *Research ideas for the classroom: Middle grades mathematics* (pp. 159-178). NY: Macmillan.

D'Zurilla, T.J., Nezu, A.M., & Maydeu-Olivares, A. (2002). *Social problem-solving inventory-revised (SPSI-R)*. North Tonawanda, NY: Multi-Health Systems, Inc.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. Resnick (Ed.), *The nature of intelligence* (pp. 231-236). Hillsdale, NJ: Erlbaum.

Harel, G. (2007). The DNR system as a conceptual framework for curriculum development and instruction. In R. Lesh, J. Kaput, E. Hamilton (Eds.), *Foundations for the future in mathematics education*. Mahwah, NJ: Lawrence Erlbaum Associates.

Harel, G. (2008). What is Mathematics? A pedagogical answer to a philosophical question. In R. B. Gold & R. Simons (Eds.), *Current issues in the philosophy of mathematics from the perspective of mathematicians*. Mathematical American Association.

Kagan, J., Rosman, B. L., Day, D., Albert, J., & Phillips, W. (1964). Information processing in the child: Significance of analytic and reflective attitudes. *Psychological Monographs, 78* (1, Whole No. 578).

Lim, K. H. (2008a). *Students' mental acts of anticipating: Foreseeing and predicting while solving problems involving algebraic inequalities and equations*. Germany: VDM Verlag.

Lim, K. H. (2008b). Mathematical knowledge for pre-service teachers. In L. D. Miller & S. R. Saunders (Eds.), *Proceedings of the US-Sino Workshop on Mathematics and Science Education: Common Priorities that Promote Collaborative Research* (pp. 92-98). Murfreesboro, Tennessee.

Nietfeld, J., & Bosma, A. (2003). Examining the self-regulation of impulsive and reflective response styles on academic tasks. *Journal of Research in Personality, 23,* 118-140.

Nygren, T. E. (2000). *Development of a measure of decision making styles to predict performance in a dynamic judgment decision making task*. Paper presented at the Annual Meeting of the Psychonomic Society, New Orleans, LA.

Van Dooren, W., De Bock, D., Hessels, A., Janssens, D., & Vershaffel, L. (2005). Not everything is proportional: Effects of age and problem type on propensities for overgeneralization. *Cognition and Instruction, 23* (1), 57–86

Schoenfeld, A. (1985). *Mathematical problem solving*. New York: Academic Press.

Schoenfeld, A. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 334-370). New York: Macmillan.

Watson, A., & Mason, J. (2007). Taken-as-shared: A review of common assumptions about mathematical tasks in teacher education. *Journal of Mathematics Teacher Education, 10,* 205-215.