July, 2012

Assessing impulsive-analytic disposition: The Likelihood-to-Act survey and other instruments

Kien H Lim, *University of Texas at El Paso*
Amy Wagler, *University of Texas at El Paso*

ASSESSING IMPULSIVE-ANALYTIC DISPOSITION: THE LIKELIHOOD-TO-ACT SURVEY AND OTHER INSTRUMENTS

Kien H. Lim — University of Texas at El Paso

Amy Wagler — University of Texas at El Paso

The likelihood-to-act (LtA) survey is a 32-item instrument that measures impulsive and analytic dispositions in solving math problems. In this research report, we compare it to other instruments related to the impulsive-analytic construct such as Frederick's Cognitive Reflection Test (CRT) and the Barratt Impulsive Scale in terms of mean scores, Cronbach alpha values, and correlation values. Both LtA-Impulsive and LtA-Analytic subscales have acceptable reliabilities of 0.79 and 0.83 respectively. The LtA-Analytic and LtA-Difference (analytic-impulsive difference) correlated well with other the Need for Cognition subscale and CRT scores. The correlations involving LtA-Impulsive subscale were unexpected and call for further investigation.

INTRODUCTION

The behaviour of "doing whatever first comes to mind … or diving into the first approach that comes to mind" (Watson & Mason, 2007, p. 307) is quite common among students while solving problems in mathematics. The term *impulsive disposition* refers to a tendency to proceed with an action that comes to mind without analysing the problem situation and without considering the relevance of the anticipated action to the problem situation (Lim, Tchoshanov, & Morera, 2009). In contrast, the term analytic disposition refers to a tendency to study the problem situation prior to taking actions. The premise underlying our work is the view that learning opportunities should be provided to help students progress from impulsive disposition to analytic disposition.

Related Theoretical Constructs

Various researchers in cognitive psychology have posited two distinct cognitive systems of reasoning: implicit-explicit (Reber, 1993), associative and rule-based (Sloman, 1996), and System 1 and System 2 (Stanovich & West, 2000). According to Evans (2006), "System 1 processes are rapid, parallel and automatic in nature: only their final product is posted in consciousness" whereas "System 2 thinking is slow and sequential in nature and makes use of the central working memory system" (p. 454). Sloman (1996) points out that the two systems often work cooperatively despite having different goals and specializing at different kinds of tasks. At times, they may each try to generate a response. A response is considered impulsive when System 1 hijacks one's attention, and reflective when System 2 overrides System's 1 response.

Whereas the dual system model can explain the impulsive-reflective distinction in terms of general functioning of cognitive processes, cognitive style can account for individual variability in impulsiveness. The *Matching Familiar Figures Test* was developed to assess children's *cognitive tempo* (Kagan et al., 1964). 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.

In mathematics education characterizing students' impulsivity-reflectivity in terms of problem-solving disposition is arguably more useful than in terms of cognitive style. A disposition is context-dependent whereas a cognitive style is a personality trait that is stable across situation and across time. In addition, viewing impulsivity-reflectivity as a continuum, rather than as a dichotomy, is more likely to influence educators to help learners progress from impulsive disposition to analytic disposition.

Instruments for Assessing Impulsive-Analytic Disposition or Related Constructs

A reliable way to investigate students' problem-solving disposition is through task-based interviews (Clement, 2000; Goldin, 1998) and think-aloud protocols (Ericsson & Simon, 1993). Impulsive anticipation and analytic anticipation can be identified from the careful analysis of students' responses to interview tasks (see Lim, 2008). Although well-suited for uncovering problem-solving disposition in individual students, this mode of data collection is not practical for large-scale assessment.

Well-designed mathematical problems can be an effective and efficient means to assessing impulsive-analytic disposition. Frederick (2005) designed a three-item test for assessing *cognitive reflection*—" the ability or disposition to resist reporting the response that first comes to mind" (p. 35). For the bat-and-ball problem in Figure 1, the wrong answer of 10 cents is considered *impulsive*. A person with only a moment of reflection would realize that the difference between $1.00 for bat and 10 cents for ball is not $1.00.

1. A bat and a ball cost $1.10. The bat costs $1.00 more than the ball. How much does the ball cost?
   Answer: ________ cents

2. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? Answer: ________ minutes

3. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?
   Answer: ________ days

Figure 1: The three items in Frederick's Cognitive Reflection Test

Another way to measure is to use questionnaire. The *Barratt Impulsiveness Scale Version 11* (BIS-11) is a 30-item self-report instrument for assessing one's impulsivity. Patton, Stanford, ad Barrrett (1995) performed a principal components analysis and confirmed three subtraits of impulsiveness: attentional, motor, and non-planning. A representative item for each subtrait is listed accordingly in Figure 2.

6. I have racing thoughts.
17. I act on "impulse".
27. I am more interested in the present than the future.

Figure 2: The three items in BIS-11

A questionnaire that assess a different but related construct is the *Need for Cognition* (NfC) scale. This 18-item self-report instrument measures one's "tendency to engage in and enjoy thinking" (Cacioppo & Petty, 1982, p. 116). An example of NfC item: "I would prefer complex to simple problem".

In the context of mathematics problem solving, Lim et al. (2009) developed the Likelihood-to-Act (LtA) survey as a means to measure impulsive-analytic disposition. For each LtA item (see Figure 3) respondents indicate on a scale of 1 to 6 how likely they are to respond to a given mathematical problem in the described manner.

1.	$78 + 987x + 654 + 321x = x + 987x + 654 + 321x$ When asked to solve for x, how likely are you to begin by studying the equation and noticing the solution?	**1**	**2**	**3**	**4**	**5**	**6**
4.	When solving a problem in mathematics, how likely are you to read and understand the problem thoroughly before deciding what to do?	**1**	**2**	**3**	**4**	**5**	**6**
17.	$90 + 1234n + 567 + 89n = n + 1234n + 567 + 89n$ When asked to solve for n, how likely are you to begin by combining like terms?	**1**	**2**	**3**	**4**	**5**	**6**
20.	When solving a problem in mathematics, how likely are you to use the first idea that comes to mind?	**1**	**2**	**3**	**4**	**5**	**6**

Figure 3: Four items in the current of the LtA instrument

The first version with nine pairs of items (one impulsive and one analytic per pair) was administered to 318 undergraduates, mostly pre-service teachers; the reliabilities of the impulsive and analytic subscales are found to be 0.64 and 0.63 respectively (Lim et al. 2009). The second version with 16 item pairs was administered to 119 pre-service and in-service teachers; the reliabilities for the two subscales are 0.74 and 0.81 (Lim & Morera, 2011). The written work of 92 participants for 6 open-ended math problems were analysed and coded; the coded scores for written responses were found to significantly correlated to both the LtA subscales (Lim & Mendoza, 2010). Based on the findings, one pair of items was replaced and seven items in the second version were modified to produce the current version of the LtA.

The 32-item version has a 2×4×4 structure: 2 types (impulsive and analytic), 4 categories (algebra, fraction, word problem, and non-mathematically-specific), and 4 items per type per category. Item 1 is analytic-algebra, Item 8 is analytic-nonspecific, Item 17 is impulsive-algebra, and Item 24 is impulsive-nonspecific.

The purpose of this research report is to (a) present students' responses to the LtA items, (b) compare the reliabilities of the LtA subscales to those of other instruments, and (b) present the correlations among measures obtained using these instruments.

METHOD

A total of 495 undergraduates, mostly pre-service teachers, participated in this study. A convenience sample involving 17 classes was used because some instructors chose not to give up their class time for us to collect data. Out of 470 participants who specified their program, 29 majored in either math or engineering, 80 are pre-service 4-8 teachers specializing in either math or science, 54 are pre-service 4-8 generalists, and the remaining 307 are pre-service elementary or bilingual or special education teachers. Out of 466 who specified their gender, 72 are males.

Within a 100-minute class period, participants took a set of three surveys (NfC, BIS-11, and LtA), took a version of a math test consisting of eight multiple-choice items (see Figure 4 for examples) and three cognitive-reflection items (see Figure 1), received warning that some of the items in the math test are "tricky", and took another version with structurally-equivalent items. Because of page limit constraint this research report does not focus on the effect-of-warning part of the study. Instead, it focuses on comparing students' scores as measured using the various instruments.

4. Benito needs to increase the amount of money he has now by 20% so that he can buy a $540 laptop. How much more money does he need?
   (a) $90
   (b) $108
   (c) $432
   (d) $2700

5. A 20-member choir takes 15 minutes to sing a song. How long does it take a 100-member choir to sing the same song?
   (a) 3 minutes
   (b) 15 minutes
   (c) 75 minutes
   (d) 95 minutes

Figure 4: Two multiple-choice items in the math test

The data analysis is based on a sample of 460 participants, with the exclusion of 23 students who had taken the LtA survey before and 12 students who had more than 2 missing entries in the LtA survey. These 460 students took an average of 8.8 minutes to complete the LtA survey, ranging from 3 to 19 minutes (based on students' self-record of start time and end time).

RESULTS AND DISCUSSION

Table 1 shows the percentage of responses for the four items in Figure 3 across the six likelihood numbers, the mean likelihood score (1 = extremely unlikely; 6 = extremely likely), and the standard deviation. For the pair of algebra items, students tended to choose higher likelihood for the impulsive Item 17 than the analytic Item 1. When comparing the mean response for Item 17 to Item 1, the estimated difference is -0.75 ($p < 0.001$). The analytic-impulsive difference is also negative for ten other math-specific pairs. As a group, the mean of the 12 math-specific analytic items is 3.61 and the mean of the 12 math-specific impulsive items is 4.64, the analytic-impulsive difference is -1.0 ($p < 0.001$). For the pair of non-math-specific items (#4 and #20), the analytic-impulsive estimated difference is 0.72 ($p < 0.001$). As a group, the mean of the four non-math-specific analytic items is 4.52 and the mean of the four corresponding impulsive items is 4.29; the analytic-impulsive difference is 0.23 ($p < 0.001$). These results suggest that students tended to think that they are analytic but when specific mathematical situations are used they tended to respond in an impulsive manner.

	Percentage of Responses						Mean Score	Std. Dev.
	1	2	3	4	5	6		
Item 1	3	6	8	22	33	28	4.59	1.32
Item 4	2	2	8	21	26	41	4.91	1.19
Item 17	2	2	3	7	25	61	5.34	1.08
Item 20	4	9	15	22	31	17	4.19	1.38

Table 1: Results for the two pairs of LtA items in Figure 3

Table 2 presents the statistics for the various instruments. LtA-Difference refers to the difference between the analytic item and the impulsive item in each pair (a negative value means more impulsive than analytic). The reliability of all the measures, except LtA-Difference, are greater than 0.7, above which is considered acceptable. The LtA-Difference has a lower reliability than the individual subscale reliabilities (0.79 and 0.83) because of the combined variability in two subscales.

Reliabilities	Number of Items	Number of Subjects	Mean	Std. Dev.	Cronbach's Alpha
LtA-Impulsive	16	460	4.55	0.69	0.790
LtA-Analytic	16	460	3.83	0.76	0.826
LtA-Difference	16	460	-0.72	0.81	0.681
BIS-11	30	458	2.02	0.33	0.804
Need for Cognition	18	459	3.30	0.57	0.817
Math-MCQ	16	426	6.83	3.20	0.738
Math-CRT	6	426	0.67	1.37	0.818
Math-Confidence	22	426	4.03	0.55	0.900

Table 2: Statistics for the measures obtained in the various instruments

Math-CRT is computed by summing the scores for the cognitive-reflection items; students averaged 0.67 corrects out of six items. When restricted to the three original CRT items (Figure 1), our students, mostly prospective teachers, averaged 0.32 out of 3 items, which is lower than those found by Fredericks (2005), ranging from 0.57 in the University of Toledo to 2.18 in MIT. In fact, 342 out of 433 students who attempted the CRT items got all three items wrong. The skewness of CRT towards zero raises the issue of its discriminatory power, especially for the population of pre-service teachers.

Math-MCQ is computed by summing the scores for the multiple-choice items in the math test; students averaged 6.8 corrects out of 16 items with a standard deviation of 3.2. Unlike Math-CRT, Math-MCQ is less skewed and more normally distributed.

For each math item, students indicated their level of confidence on a 5-point scale (1="I'm certain I'm wrong, 2="I think I'm wrong" … 5="I'm certain I'm right"). Math-Confidence is computed by averaging their confidence scores across the 22 items. Interestingly, students generally thought that they are correct (mean of 4.03) yet got only got 43% and 11% correct for the MCQ items and CRT items respectively. More specifically, 12% and 57% of students got Item 4 and Item 5 in Figure 4 correct, and only 10%, 12%, and 10% for the three CRT items in Figure 1. These results are consistent with Frederick's (2005) findings that respondents who missed the problems thought they were easier than the respondents who got them right.

Table 3 presents the Pearson correlations and associated false discovery rate adjusted *p*-values among the mean scores (Benjamini & Yekutieli, 2001). The highlighted correlations indicate that the LtA-Impulsive did not behave as expected. LtA-Impulsive items should be in oppositon to LtA-Analytic items, yet the mean scores exhibit significant correlations with the LtA-Analytic, with the NfC, and negatively with the BIS-11 mean scores. On the other hand, the significant negative correlation between LtA-Impulsive and Math-CRT mean scores suggests that LtA-Impulsive items might be valid in assessing non-cognitive reflection.

	LtA-I	LtA-A	LtA-D	BIS	NfC	MCQ	CRT	Conf
LtA-Impulsive	1							
LtA-Analytic	0.38**	1						
LtA-Difference	-0.49**	0.62**	1					
BIS-11	-0.21**	-0.16**	0.03	1				
NfC	0.21**	0.37**	0.17**	-0.45**	1			
Math-MCQ	-0.04	0.09	0.12*	-0.12*	0.25**	1		
Math-CRT	-0.16**	0.20**	0.34**	-0.07	0.24**	0.51**	1	
Math-Confidence	0.18**	0.28**	0.12*	-0.22**	0.28**	0.28**	0.31**	1

Table 3: Correlations among various scores (** $p<0.01$; * $p<0.05$)

The LtA-Difference scores are behaving as expected since they are significantly correlated with the other measures except BIS-11, which measures impulsivity in everyday actions rather than in mathematical situations. This may explain why the negative correlation between BIS-11 and Math-CRT is not significant.

Interestingly, BIS-11 and Math-Confidence means scores are significantly negatively correlated, which suggests students who are impulsive in life tend to be not confident in math. The correlation of 0.24 between NfC and Math-CRT is consistent with the correlation of 0.22 found in Frederick's (2005) study.

The high correlation between Math-CRT and Math-MCQ can be taken to mean that the two set of math items are related by a common factor—possibly the impulsive-analytic disposition. Both Math-CRT and Math-MCQ correlated equally well with NfC. The LtA scales correlate better with the Math-CRT than with the Math-MCQ.

CONCLUSION

The LtA instrument is still in its early phases of development. By having four categories of items in the LtA, we learn that non-math specific items seem to behave differently from the other three math-specific categories. This finding is consistent with the results we found in our principal factor analysis, which will be presented elsewhere. In this paper, we investigate the criterion-related validity for the LtA survey by examining the Pearson correlations between the LtA-Impulsive, LtA-Analytic, LtA-Difference, BIS-11, NfC, Math-MCQ, Math-CRT, and Math-Confidence mean scores. Results support the criterion-related validity of the LtA-analytic subscale and LtA-Difference. The unexpected correlations involving the LtA impulsive subscale call for further investigation. We are currently analysing interview data to study 15 respondents' problem-solving dispositions and their interpretations of the LtA items.

In addition, our study confirms reliability of three established instruments. Our results involving pre-service teacher corroborate with those involving undergraduates reported by Frederick (2005); our study adds credence to his three-item CRT. The strong correlations between NfC score and other scores, except LtA-Impulsive, suggest students' high need for cognition is related to analytic disposition and cognitive reflection. We also found that the type of impulsiveness (attentional, motor, and non-planning), as measured by BIS-11, does not seem to be tightly coupled to mathematical impulsiveness, as measured by CRT and LtA-Difference.

References

Ben-Zeev, T., & Star, J. R. (2001). Spurious correlations in mathematical thinking. *Cognition and Instruction, 19,* 253–275.

Benjamini, Yoav; Yekutieli, Daniel (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of* Statistics, 29, 1165–1188.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42*, 116-131.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (Revised Ed.). Cambridge, MA: MIT Press.

Evans, J. St. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Science, 7*, 454-459.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*(4), 25-42.

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

Lim, K. H., & Mendoza, M. (2010). Inferring impulsive-analytic disposition from written responses. In P. Brosnan, D. B. Erchick, D. B. & L. Flevares, L. (Eds.), *Proc. 32nd Conf. of the North America Chapter of the Int. Group for the Psychology of Mathematics Education* (pp. 759-767). Columbus: Ohio State University.

Lim, K. H., & Morera, O. (2011). Impulsive-analytic disposition: Instrument pilot testing. In L. R. Wiest & T. Lamberg (Eds.), *Proc. 33rd Conf. of the North America Chapter of the Int. Group for the Psychology of Mathematics Education* (pp. 284-292). Reno, NV: University of Nevada.

Lim, K. H., Morera, O., & Tchoshanov, M. (2009). Assessing problem-solving dispositions: Likelihood-to-act survey. In S.L. Swars, D.W. Stinson & S. Lemons-Smith (Eds.). *Proc. 31st Conf. of the North America Chapter of the Int. Group for the Psychology of Mathematics Education* (pp. 700-708). Atlanta: Georgia.

National Research Council. (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academic Press.

Patton, J. M., Stanford, M. S., & Barratt E. S. (1995). Factor structure of the Barratt Impulsiveness Scale. *Journal of Clinical Psychology, 51*, 768-774.

Radatz, H. (1979). Error analysis in mathematics education. *Journal for Research in Mathematics Education*, 10, 163-172.

Reber, A. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York: Oxford University Press.

Sloman, S. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin 119*, 3-22.

Stanovich, K., & West, R. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioural and Brain Sciences 23*, 645-726.

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.