July, 2012

Impulsive-analytic disposition in mathematical problem solving: A survey and a mathematics test

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

Impulsive-Analytic Disposition in Mathematical Problem Solving: A Survey and a Mathematics Test

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

Amy Wagler
University of Texas at El Paso
< awagler2@utep.edu>

The Likelihood-to-Act (LtA) survey and a mathematics test were used in this study to assess students' impulsive-analytic disposition in the context of mathematical problem solving. The results obtained from these two instruments were compared to those obtained using two widely-used scales: Need for Cognition (NFC) and Barratt Impulsivity Scale (BIS). The exhibited correlations of the LtA scores with the NFC, BIS, and a math test provide evidence of the criterion validity of the analytic LtA items, and suggests further revision of the impulsive LtA items to improve the overall measurement validity of the LtA scale. Students LtA scores were found to be marginally correlated to their math scores and correlated to their confidence levels in the math items.

Teachers often see students solve mathematical problems using familiar tools or strategies without seemingly to analyse the problem situation. Watson and Mason (2007) describe this as "doing whatever first comes to mind … or diving into the first approach that comes to mind" (p. 307). Lim, Tchoshanov, and Morera (2009) started developing an instrument, called the Likelihood-to-Act survey, to assess students' impulsive-analytic disposition. *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. *Analytic disposition*, on the other hand, refers to a tendency to study the problem situation prior to taking actions. In this paper, scores of students' impulsivity obtained using a recent version of the Likelihood-to-Act survey and a two-version mathematics test are presented and are compared to two other scores obtained using established instruments in the field of psychology.

Theoretical Background

Impulsivity and reflectivity have been contrasted in terms of cognitive tempo or response style. Kagan, Rosman, Day, Albert, and Phillips (1964) constructed the Matching Familiar Figures Test to measure children's cognitive tempo. An impulsive child is one with response time being faster than the median and accuracy rate being lower than the median, whereas a reflective child is one with response time being slower and accuracy rate being higher than the median. In a study on consistency in cognitive responses among adults across academic tasks, Nietfeld and Bosma (2003) 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. In our study, mathematical tasks are designed to assess whether students respond to the first idea that comes to mind or whether they analyse the problem situation.

Impulsive and analytic responses to a situation can be accounted using the *dual process theories* (Evans, 2003; Stanovich & West, 2000) from cognitive psychology. The basic tenet in these theories is that two modes of cognitive processing are at work. The "intuitive" mode has these characteristics: automatic, fast, preconscious, low effort, associative, parallel processing, and little variation across individuals. The "analytical" mode has the opposite characteristics: controlled, slow, conscious, high effort, rule-based, serial processing, and greater variation across individuals (Frankish, 2010). Dual process theories are gaining attention in mathematics education recently. Gillard, Van Dooren, Schaeken, & Verschaffel (2009) used them to account for why people fail to solve mathematical tasks they should be able to solve correctly given their mathematical knowledge and skills. Leron (2010) suggested a bridge between intuitive and analytic thinking can deepen student conceptual understanding. We, on the other hand, focus on creating teacher awareness about students' disposition in solving mathematics problems. This effort initiated the development and testing of a survey instrument.

Instruments for Assessing Constructs related to Impulsiveness

The *Need for Cognition* (NFC) scale is a self-report instrument for measuring one's "tendency to engage in and enjoy thinking" (Cacioppo & Petty, 1982, p. 116). This instrument consists of 18 Likert-scale items. Presented below are three sample items:

- NFC_6: I find satisfaction in deliberating hard and for long hours.
- NFC_7R: I only think as hard as I have to.
- NFC_14: The notion of thinking abstractly is appealing to me.

The *Barratt Impulsiveness Scale Version 11* (BIS) is a 30-item self-report instrument for assessing impulsivity in one's behaviours. Patton, Stanford, and Barrrett (1995) performed a principal components analysis and found six first order factors. A representative item for each factor is presented below.

- BIS_5: I don't "pay attention." (Attentional)
- BIS_6: I have "racing" thoughts. (Cognitive Instability)
- BIS_2: I do things without thinking. (Motor)
- BIS_16: I change jobs. (Perseverance)
- BIS_12R: I am a careful thinker. (Self-control)
- BIS_10R: I save regularly. (Cognitive Complexity).

The current version of the *Likelihood-to-Act* (LtA) survey has 32 items which can be sub-divided into 16 pairs of items, with one impulsive item and one analytic item in each pair. The 16 pairs can be divided equally into four subcategories: algebraic, proportional, fraction, and general. The first four items in the instrument, and their corresponding counterparts, are presented in Figure 1. For example, LtA_1 is an analytic-algebraic item and LtA_17 is an impulsive-algebraic item.

The data reported in this paper is part of a larger project that seeks to develop, test, and refine the LtA instrument. The first version of LtA has nine pairs and was administered to 318 undergraduates, mostly pre-service teachers; the reliabilities of the impulsive and analytic subscales were 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 of 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 reliabilities of the two subscales in the current version are 0.79 and 0.83. In this paper, the criterion validity of the LtA instrument is investigated by inspecting the correlations between the four subcategories of the LtA instrument and the two established scales: NFC and BIS.

For each item, indicate as honestly as you can how likely you are to act in the manner specified in the statement by circling the number using this scale:

1 = Extremely Unlikely
2 = Unlikely
3 = Somewhat Unlikely
4 = Somewhat Likely
5 = Likely
6 = Extremely Likely

Start Time:
:

Item						
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
2. Given that 4 candy bars cost \$2.15. When asked to find the cost of 40 candy bars, how likely are you to begin by EITHER finding the cost per candy bar OR setting up a proportion?	1	2	3	4	5	6
3. $\frac{3}{4} + \frac{1}{10} + \frac{9}{10}$ When asked to find the answer for the above arithmetic expression without using a calculator, how likely are you to begin by finding the common denominator?	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
18. Given that 6 bottles of mineral water cost \$2.10. When asked to find the cost of 30 bottles, how likely are you to notice the relationship between 30 and 6 and use it use to obtain the answer?	1	2	3	4	5	6
19. $\frac{3}{5} + \frac{11}{12} + \frac{1}{12}$ When asked to find the answer for the above arithmetic expression without using a calculator, how likely are you to begin by studying the fractions to see if you can predict the answer?	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 1. Sample items in the LtA instrument

In this paper, the results obtained from a mathematics test that contain items designed to elicit impulsive responses are included in the analysis. Frederick (2005) developed a three-item *Cognitive Reflection Test* (see Figure 2) to assess one's "ability or disposition to resist reporting the response that first comes to mind" (p. 35). We expanded his test by including eight multiple-choice items (see Figure 3). For each of the 11 items, a "1" is assigned if students gave an "impulsive" response (e.g. 10¢ for Item 9). Students were expected to indicate their confidence level. We created two structurally-equivalent versions (Version Y and Version Z) to investigate whether a warning that some of the items were "tricky" would affect students' performance in the second version. Because of page limit constraint this paper does not focus on the effect-of-warning part of the study.

9. 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

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

11. 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 2. The three Cognition Reflection Test items in the Math Test – Version Y

3. Captain Toot was 27 years old when he had 8 years of sailing experience. How old was he when he had 16 years of sailing experience?
   (a) 35 years old
   (b) 51 years old
   (c) 54 years old
   (d) 81 years old

   ☐ I'm certain I'm right
   ☐ I think I'm right
   ☐ I've a 50-50 chance of being right
   ☐ I think I'm wrong
   ☐ I'm certain I'm wrong

6. The ratio of girls to boys in a Science club at Mesa High School is 2:5. There are 35 students in the Science club. How many members are girls?
   (a) 10
   (b) 14
   (c) 15
   (d) 25

   ☐ I'm certain I'm right
   ☐ I think I'm right
   ☐ I've a 50-50 chance of being right
   ☐ I think I'm wrong
   ☐ I'm certain I'm wrong

Figure 3. Two multiple-choice items in the Math Test – Version Y

Research Method

A total of 495 undergraduates, mostly pre-service teachers, participated in this study. A convenience sample involving 17 classes was used. 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, and LtA), took a version of the 11-item math test, received warning about the items being "tricky", and took the second version of the 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.

Results

The NFC and BIS scales were evaluated for construct validity by fitting confirmatory factor models to the items. All items that load to a single factor were retained for further analysis. For the NFC scale, only items 6, 8, and 18 do not load to a single factor and are eliminated from further analysis. For the BIS scale, 19 items (3, 4, 6, 7, 11, 15, 16, 18, 19, and 21-30) do not load and were eliminated from further analysis. The Pearson correlation between the NFC and BIS sum scores is -0.442 (p-value is <0.0001). The negative correlation between NFC and BIS is expected in that students who enjoy thinking tend to be less impulsive.

Table 1 shows students' mean scores for each subcategory of the LtA instrument. Students' impulsive mean scores are higher than analytic mean scores for the first three subcategories which contain mathematically-specific items. The impulsive mean scores, on the other hand, are lower for the general subcategories. The impulsive mean scores are found to be significantly different from the analytic mean scores for all four subcategories.

Table 1
Results for the Four Subcategories of LtA Items

Subcategories	Impulsive	Analytic	Test Statistics	*p*-value
Algebraic	4.76	3.57	17.42	<0.0001
Fraction	4.84	3.41	20.93	<0.0001
Proportional	4.33	3.83	8.58	<0.0001
General	4.29	4.52	-3.76	0.0002

Recall that the LtA items are paired so as to measure impulsivity along an impulsive-analytic continuum. Thus, the differences (impulsive minus analytic) could be computed to assess students' impulsivity along this continuum. The impulsive mean scores, analytic mean scores, and impulsive-analytic difference scores for the four subcategories were analysed for any relationship with the NFC and BIS sum scores. Table 2 show the results of the estimated correlations and *p*-values, adjusted for maintaining an overall 5% error rate for the family of inferences using the Hochberg (1988) procedure.

- The analytic LtA scores, for all four subcategories, are positively correlated with the NFC scores and are somewhat negatively correlated with the BIS scores.
- The impulsive LtA scores demonstrate less association with the NFC and BIS sum scores. Interestingly, the NFC sum scores have significant positive correlation with impulsive-proportional and impulsive-general subcategories. The BIS sum score has significant negative correlation with impulsive-general subcategory.
- The impulsive-analytic difference scores are all strongly correlated with the NFC sum score for three subcategories, excluding the proportion subcategory. None of the four subcategories are correlated with the BIS sum scores.

Table 2

Estimated Pearson Correlations between LtA Sub-scores and the Other Two Survey Scores

	Need for Cognition		Barratt Impulsiveness Scale	
Subcategories	Pearson Corr.	*p*-value*	Pearson Corr.	*p*-value*
Analytic Items				
Algebraic	0.388	<0.0001	-0.225	0.0011
Fraction	0.291	<0.0001	-0.148	0.0398
Proportional	0.286	<0.0001	-0.115	0.0716
General	0.429	<0.0001	-0.294	<0.0001
Impulsive				
Algebraic	0.118	0.1882	-0.157	0.0669
Fraction	0.072	0.2626	-0.071	0.2626
Proportional	0.283	<0.0001	-0.137	0.1232
General	0.217	0.0040	-0.192	0.0140
Impulsive-Analytic Diff.				
Algebraic	-0.216	0.0050	0.025	0.6987
Fraction	-0.181	0.0258	0.021	0.7426
Proportional	-0.015	0.8206	-0.040	0.5266
General	-0.202	0.0097	0.136	0.1642

*p-values are adjusted to control the family-wise error rate at 5% (Hochberg, 1988)

The math test items were analysed for any relationship with the NFC and BIS sum scores and the combined (16 pairs) LtA difference scores. Polychoric correlations (Olsson, 1979) are utilized since the correctness scores are not continuous but are scored with a “1” for correct response and a “0” for incorrect response, and the resulting sum scores are the number of correct responses. The polychoric correlations (for correct responses and impulsive responses) or Pearson correlations (for confidence levels) and *p*-values with the Hochberg adjustment are presented in Table 3.

Table 3

Estimated Polychoric Correlations between Scores from Math Tests and the Survey Scores

	Need for Cognition		Barratt Imp. Scale		LtA Difference	
Survey Scores	Corr.	*p-value**	Corr.	*p-value**	Corr.	*p-value**
Correct Responses						
Version Y	0.250	0.0007	-0.103	0.3190	-0.135	0.0524
Version Z	0.232	0.0021	-0.066	0.4224	-0.142	0.0503
Impulsive Responses						
Version Y	-0.225	0.0026	0.111	0.3190	0.196	0.9994
Version Z	-0.175	0.0284	0.051	0.4224	0.171	0.9973
Confidence Level						
Version Y	0.300	<0.0001	-0.229	0.0022	-0.194	0.0039
Version Z	0.235	0.0019	-0.193	0.0136	-0.131	0.0528

*p-values are adjusted to control the family-wise error rate at 5% (Hochberg, 1988)

Consistent with previous results, the NFC items are correlated with the correctness scores in the math test and negatively correlated with the impulsive-response scores for both versions. The BIS items do not exhibit strong correlations. However, like correctness scores, students' confidence level scores in the math test are positively correlated to NFC sum scores but negatively correlated to BIS sum scores. The LtA paired items are marginally correlated with the correctness score and with the confidence level ratings, but not the impulsive responses to the math test.

Discussion

The results in Table 1 suggest that students tend to respond more impulsively to criterion-specific items like algebra (#1 and #16), proportion (#2 and #17), and fraction (#3 and #18), but more analytically to items with general descriptors (#4 and #20). In terms of dual-process theories, students tend to suspend their "analytical" mode of processing in response to familiar situations (e.g. adding fractions) especially when they have automatic responses (finding common denominators) operating at the "intuitive" level.

The strong positive correlations between LtA-analytic and NFC (see Table 2) suggest that students who agreed with the analytic statements in the LtA survey also tend to agree with the subset of NFC items (excluding items 6, 8, and 18). These results suggest that the LtA analytic items are measuring the same construct as the NFC measures and hence strengthen the criterion validity of the analytic items. The negative correlation between LtA-analytic and BIS suggest that students who agree with the analytic statements in the LtA instrument tend to disagree with the subset of BIS items (1, 2, 5, 8, 9, 10, 12, 13, 14, 17, and 9). The negative correlation was not significant for the proportion subcategory. This lack of statistical significance implies that the set of proportion-related items should be further examined and revised to strengthen the criterion validity.

The correlations for LtA-impulsive items are not as significant as those for LtA-analytic. Three correlations (between proportional and NFC, between general and NFC, and between general and BIS) were significant but they were contrary to what was expected. These unexpected correlations imply that the impulsive LtA items should undergo further refinement in order to improve the criterion validity of the scores.

The correlations between LtA impulsive-analytic difference scores and NFC scores are strong except for the proportion subcategory. This result suggests that the set of paired items for assessing how student react to proportion-related items should be further examined for criterion validity. The correlations between LtA difference scores and BIS scores are not significant. Whereas NFC items are about cognitive activity, BIS items tend to assess impulsivity in a general sense. The connection between impulsive behaviours in general settings and those in educational settings may not be clear. This may explain the lack of relationship between the LtA difference scores and the BIS sum score.

The two versions of the math test have very similar results; this can be taken as an indicator of reliability of the two versions with different but structurally equivalent items. The NFC sum score show evidence of association with all scores related to the math tests. The positive correlation between the NFC score and math correctness score implies that students who score high on the NFC scale also tend to get more math items correct. Conversely, the negative correlation between the NFC and the impulsive-response score implies students with high NFC score tend to choose the impulsive answer less frequently. The positive correlation between the NFC score and the confidence-level score suggest students with high NFC score tend to report high levels of confidence associated with their answers to the math test items.

The relationship between the BIS sum scores with the confidence levels provides evidence that students with high BIS score tends to report low levels of confidence associated with their answers on the math tests. Additionally, it may be inferred from the relationship between the LtA impulsive-analytic difference scores and the math test items and the confidence levels, that students who tend toward impulsive behaviours when approaching cognitive tasks also tend to score lower on the math tests and have less confidence associated with their answers on the math tests. There is no evidence of a relationship between the LtA impulsive-analytic difference scores and the number of impulsive answers on the math tests.

References

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

Evans, J. (2003). In two minds: Dual process accounts of reasoning. *TRENDS in Cognitive Sciences, 7*(10), 454-459.

Frankish, K. (2010). Dual-process and dual-system theories of reasoning. *Philosophy Compass 5*(10), 914-926.

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

Gillard, E., Van Dooren, W., Schaeken, W., & Verschaffel, L. (2009). Dual processes in psychology of mathematics education and cognitive psychology. *Human Development, 52*(2), 95-108.

Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika, 75*, 800–803.

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

Leron, U. (2010). The power of natural thinking: Applications of cognitive psychology to mathematics education. In P. Brosnan, D. B. Erchick & L. Flevares (Eds.), *Proceedings of the 32nd annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. VI, pp. 9-20). Columbus, OH: The Ohio State University.

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.), *Proceedings of the 32nd annual meeting of the North American Chapter of the International 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. *Proceedings of the 33rd annual meeting of the North American Chapter of the International 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.). *Proceedings of the 31st annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 700-708). Atlanta: Georgia: PME-NA.

Nietfeld, J., & Bosma, A. (2003). Examining the self-regulation of impulsive and reflective response styles on academic tasks. *Journal of Research in Personality, 37*, 118–140.

Olsson, U. (1979) Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika 44*, 443-460.

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

Stanovich, K.E., & West, R.F. (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.