DOES CHRONOMETRY HAVE A PLACE IN ASSESSING MATH DISORDERS?

A. Alexander Beaujean, Andrew Knoop, and Gregory Holliday

Abstract **The purpose of this pilot study was to determine if a single math-based chronometric task could accurately discriminate between college students with and without a diagnosed math disorder. Analyzing data from 31 students (6 in the case group, 25 in the clinical comparison group), it was found that the single chronometric task could accurately predict students who did not have a diagnosed math disorder, but not students who had the diagnosis. Moreover, no other non-math psychometric task could add to the predictive power of the chronometric task, indicating that the role of chronometry warrants further study.**

A. ALEXANDER BEAUJEAN, Ph.D., is a graduate student, University of Missouri, Columbia.
ANDREW J. KNOOP, Ph.D., is clinical professor, University of Missouri, Columbia.
GREGORY A. HOLLIDAY, Ph.D., is clinical associate professor, University of Missouri, Columbia.

Scientists have long made use of mental chronometry (Jensen, 1985). Chromometric (i.e., time-based measures) tasks made their way into psychological research in the late 1800s with the work of Franciscus Donders (1868-1869/1969) and Francis Galton (1883). While Donders was interested in using chronometric tasks to uncover details involved in elemental cognitive processes (e.g., discrimination, apprehension), Galton was more concerned about using them to study individual differences. Donders' research agenda has been carried on in the field of experimental psychology (Meyer, Osman, Irwin, & Yantis, 1988), whereas Galton's branch of inquiries largely halted after his death. In the late twentieth century though, a group of researchers began to reinvestigate Galton's ideas and produced a formidable literature on the subject (Deary, 2000; Jensen, 1998; Vernon, 1987).

While a significant portion of research in the individual differences renaissance sought to determine how performance on time-based tasks related to overall cognitive functioning, defined as *g* by Spearman (1904), others sought to see how they differed across various psychopathologies (Posner, 1978). Some have even argued that chronometric tasks have the potential for use in clinical batteries (i.e., assessing the presence of learning disorders; Jensen, 1987; Kulak, 1993; Swanson, 1987). For example, Jensen (1987) and Kranzler (1994) assert that research and theory on learning disabilities (LD) cannot significantly advance until researchers and clinicians use instruments that are more sensitive to the nexus of the brain-behavior relationship than are more traditional psychometric instruments (e.g., the Wechsler scales). Consequently, they have suggested the use of mental chronometric instruments. Because chronometric tasks minimize extraneous influences on task performance (e.g., task strategies), they tend to be much more sensitive to the individual differences involved in various aspects of information processing. If one were to develop subject-specific chronometric tasks (e.g., math facts), it would be assumed that they could better measure individual differences in information processing within that specific domain than more traditional psychometric instruments.

To date, minimal research has been conducted in the field of LD with chronometric tasks, especially concerning processing differences between those with and without an LD. Existing findings have been encouraging, however. For example, Whyte, Curry, and Hale (1985) and Kranzler (1994) found significant differences in the performance of various chronometric tasks between individuals with and without a reading disorder, although the specific tasks used (e.g., inspection time) tend to be a proxy for general cognitive ability (Kranzler & Jensen, 1989; Luciano et al., 2004). Cisero, Royer, Marchant, and Jackson (1997), in a series of studies with chronometric tasks designed specifically to assess various components of the reading process, found that college students with reading learning disabilities consistently took longer to process and respond to various lexical stimuli (e.g., word and pseudo-word naming) – a result that not only differentiated them from a non-diagnosed sample, but also from a "generalized learning disability" sample. Likewise, Hayes, Hynd, and Wisenbaker (1986) found significant differences on a battery of chronometric tasks between college students with and without a diagnosed LD, even after controlling for cognitive ability. Nonetheless, Kranzler (1994) writes that while many of the studies have suffered from various methodological inadequacies, these do not allow for a disconfirmation of other competing hypotheses as to why there might be between-group differences in performance (e.g., differences in the participants' knowledge base elaborateness and structure; Ceci, 1990). Thus, there is a need for much more research in this area.

One particular area in need of more research is the place of chronometry in assessing math-based LDs (MLD). People with math difficulties are a heterogeneous group (Fleischner & Manheimer, 1997; Knoop, Beaujean, & Holliday, 2005), with only a subset of this "person space" being composed of people with an MLD. Differentiating the MLD subpopulation from other subpopulations, especially those with other learning exceptionalities, can be difficult using current psychometric techniques (Gonzalez & Espinel, 1999).

Thus, given the relative dearth of research literature on the use of chronometric tasks within the LD arena – especially concerning MLDs – coupled with the increasing need to identify and understand the specific sources of individual differences within this population, this pilot study was designed to begin the much larger process of determining the validity and utility of time-based tasks as assessment and identification tools. Specifically, the purpose of the study was to determine whether a math-based chronometric task could discriminate between individuals with and without a diagnosed MLD.

METHOD

This pilot study was designed to determine if a math-based chronometric task could be used to predict mathematics disorders as diagnosed by psychometric instruments; namely, the math subtests of the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, & Mather, 2001). It was decided that if this study found that the chronometric task had good diagnostic properties, it would be the basis for a larger, more elaborate study.

Participants

Thirty-one college students participated in this study, as part of a larger study on college students with math difficulties (see Table 1 for descriptive statistics of sample). The larger study, supported by the National Science Foundation, explored learner characteristics of college students with math difficulties through the administration of a traditional psychoeducational assessment battery. Participants were all self-referred to a campus-based psychoeducational clinic because of mathematics difficulties. The chronometric tests were administered to some of the self-referred participants as supplemental instruments to the psychoeducational battery.

Of the 31 participants, 6 met the DSM-IV-TR (American Psychiatric Association, 2000) definition of a mathematics disorder, in that they each exhibited (a) mathematical ability as measured by a standardized test below the expected range, given the individual's age, cognitive ability and education (IQ-Achievement discrepancy); (b) significant interference in academic tasks that require mathematics ability; and (c) significant mathematics difficulty not attributed to, or over and above, any sensory deficits. Additionally, none of the six individuals with an MLD met DSM-IV-TR diagnostic criteria for any other cognitive, learning or emotional disorder, thereby eliminating any statistical confounds introduced through comorbity. These participants comprised the case group. The other 25 participants met no DSM-IV-TR diagnostic criteria, and thus comprised a clinical comparison group (Kazdin, 2002).

Instrumentation

All participants completed both a psychometric and a chronometric battery. The psychometric battery consisted of (a) Wechsler Adult Intelligence Scale-Third Edition (WAIS-III; Wechsler, 1997a); (b) Wechsler Memory Scale-Third Edition (WMS-III; Wechsler, 1997b); and (c) the Mathematics and Reading subtests (six subtests total) of the Woodcock-Johnson III Tests of Achievement (WJ-III; Woodcock et al., 2001).

The chronometric battery consisted of two trials of a mathematics task administered via the Computer-Based Academic Assessment System (CAAS; Royer, 1999). The CAAS collects, in milliseconds, both the median reaction time (RT) and standard deviation reaction time (SDRT) of participants performing the mathematics task. For this study, participants took a triple multiplication subtest, which measures mathematics fluency for complex multiplication facts, such as: 6 x 4 x 2 = ___. The participant responds into a microphone, and, for correct answers, the computer registers and calculates the time between the presentation of the item on the computer screen and the verbal response provided by the participant. Two trials of 20 similar items each were administered. Both trials take a total of approximately 15 minutes to complete. Before taking the chronometric tasks, each participant completed an untimed, paper-and-pencil version of the tasks with 100% accuracy.

RESULTS

Model Selection

Participants' results from the psychometric and chronometric tasks were put into a logistic regression to see: (a) if the chronometric task could discriminate between the MLD and no diagnosis (ND) groups, and (b) if a particular combination of the administered tasks would best discriminate between the two groups. Initially, the CAAS RT variable was put into the model, and the model's deviance was obtained (see Table 2). From this initial model, other variables were added to see if they improved the model's fit. If adding other variables to the model resulted in a better fit than the original model (CAAS RT only), the change in deviance (Δ Deviance) would be statistically significant (Neter, Kutner, Nachtsheim, & Wasserman, 1996).

The results of the model selection process revealed that no other variable significantly improved the model's fit, suggesting that no other subtest in the data set provided a significant improvement in the ability to discriminate between the MLD and ND groups beyond that of the CAAS RT (see Table 1 for mean values of the psychometric and chronometric tasks).

Model Fit

Once the variables for the model were ascertained, the overall model was assessed for fit. As most formal fit

Table 1
Descriptive Statistics of Groups

	Comparison	Case
N	25	6
Age	25.32	22.33
Race[1]	84	100
Sex[2]	68	100
ACT[5]	22.44[3]	22.50[4]
WAIS-III FSIQ[6]	111.96	108.17
WMS-III GM[7]	107.48	106.33
WJ-III TA[8]	107.40	107.00
CAAS RT[9]	6.10	9.972

1. % Caucasian. 10% were African American and 3% were Hispanic in the comparison group.
2. % Female.
3. *n* =18.
4. *n* =4.
5. Average ACT score.
6. Average full scale IQ score on WAIS-III.
7. Average general memory score on WMS-III.
8. Average total achievement score on WJ-III.
9. Average reaction time on CAAS task.

Table 2
Fitted Models

Model	Variables	Deviance	Δ Deviance	*df*	*p*
1	CAAS RT	16.044			
2	CAAS: RT & RTSD	14.952	1.092	1	0.296
3	CAAS RT, FSIQ	16.015	0.029	1	0.865
4	CAAS, WMS General Memory	15.819	0.225	1	0.635
5	CAAS RT, WJ-III Total Achievement	16.040	0.004	1	0.950

Note. Deviance equals -2 Log Likelihood of the variables, so the change in Deviances (Δ Deviance) are distributed as chi-squares.

indices require a large *n*, this study employed two informal fit indices. First, an appropriate model must demonstrate that "the estimated response function for the data is monotonic and sigmoidal in shape" (Neter et al., 1996, p. 590). To do this, a logit value was generated for each participant.[1] Then, the sample was rank ordered and quadripartitioned into approximately equal-sized groups based on the rank order of the logit values (i.e., the six lowest values, the next six lowest logit values, etc.). For each group, an average logit value was calculated, as well as the proportion of the participants who belonged to the MLD group.

The average logit values and MLD proportion were graphed (see Figure 1). The resulting curve is both monotonic and approaching a sigmoidal shape, providing evidence that the selected model was appropriate.

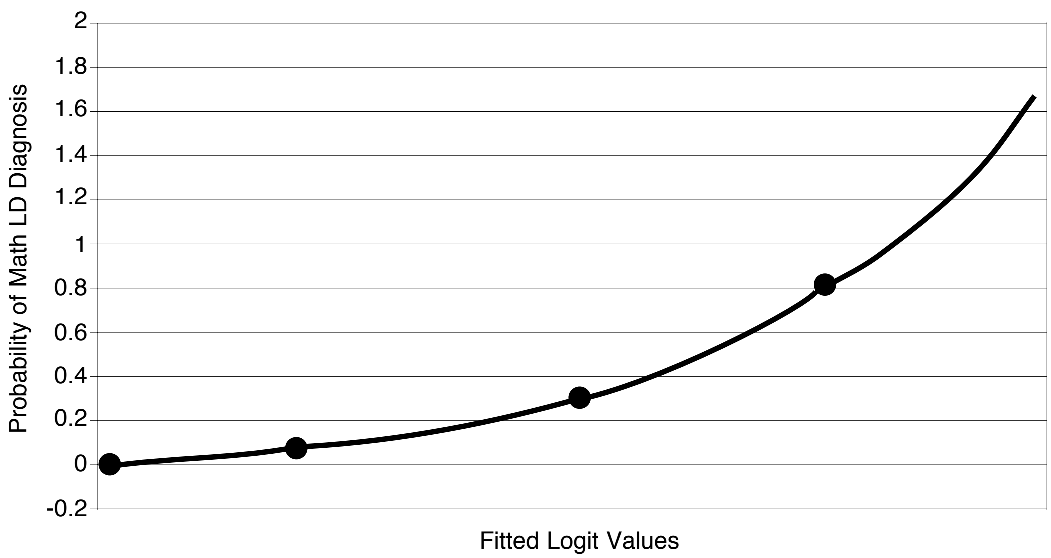
Second, a 2x2 table was generated to assess the percentage of correct and incorrect classifications based on each participant's probability of an MLD (see Table 3).[2] Data from Table 3 indicate that this model

Table 3
Classification Table Using Just CAAS RT

		Predicted		% Correct
		No MD	MD	
Actual	No MD	24	1	96
	MD	3	3	50
Overall Percentage				87.1

Note. If individual i had a probability score greater than one half (i.e., $\pi_i > .5$), he/she was classified in the MD group.
Odds ratio for this model: 4.616.

Figure 1. Median logit value by probability of diagnosis for the quadripartitioned sample.



predicted members of the ND group with approximately 96% accuracy. However, the model was less accurate (50% accuracy) when predicting group membership for the MLD group. Thus, it would appear that reaction time on this single chronometric task is specific (i.e., ability to correctly classify the comparison group), but not very sensitive (i.e., ability to correctly classify the case group).

CONCLUSION

Overview of Results

The purpose of this study was to determine if chronometric tasks could be used in a psychoeducational assessment battery to aid in the determination of an MLD. To that end, the study assessed the predictive capability of the median reaction time on a computer-based math task to discriminate between groups of collegiate participants with and without a diagnosed MLD. It was found that the chronometric task was able to predict those who did not have a diagnosed MLD better than chance (96% accuracy), but was not able to predict those who had an MLD, thus suggesting the task is specific but not very sensitive.

These results are encouraging for the future use of chronometry in clinical assessment. Because of its pilot nature, the study used only one math-based chronometric task and included a sample comprised of case and clinical comparison groups. Both these aspects of the study were not favorable for determining whether the chronometric task could discriminate between groups, yet performance on the chronometric task was able to successfully predict those who did not have a diagnosed MLD with approximately 96% accuracy.[3] Moreover, the addition of any psychometric tasks to the results from the chronometric tasks did not significantly improve the model's fit or prediction ability.

Implications for Practice

The results of this study can help in developing the future educational use of time-based, content-specific tasks related to learning disorders. A triple multiplication task was used with college students in this study, but more basic tasks may be designed to measure acquired mathematical skills and concepts in much younger children. For example, elementary-aged students, especially those identified as being at risk for learning difficulties, could be assessed on such elementary, time-based tasks as number naming, simple addition, and simple subtraction facts, and in later years, multiplication and division facts. This close monitoring of at-risk students as they assimilate increasingly complex arithmetical concepts in the early stages increases the chances of earlier identification and remediation (Geary, 1994).

Chronometric tasks are relatively easy to administer (i.e., they are computer-based), take a short amount of time to complete (the two trials of the task used in this study took less than 20 minutes total), and do not require significant time periods between repeated administrations. Moreover, specific chronometric tasks can be administered to the same student frequently during a semester or grading period. These attributes are not true of most traditional psychometric instruments, suggesting that chronometry has the potential of holding a unique place in psychoeducational assessment as either supplemental or alternative diagnostic tools; alternatively, because they showed good specificity, chronometric tasks might be best used as screening instruments.

This chronometric methodology may also be used in conjunction with curriculum-based measurement (CBM). In CBM, teachers design brief, simple tests (probes) to assess how well students are learning basic skills currently being taught in the classroom. These simple probes can be quickly administered over time to track content mastery. CBM probes used in conjunction with chronometric technology will increase CBM's efficacy as a screening tool for early academic concerns. Teachers would then be able to measure a student's content mastery with regard to current classroom concepts, and concurrently, through analysis of chronometric data, be able to screen the same student for underlying information-processing differences that may signal current or future academic difficulties. School psychologists or other qualified professionals, equipped with common technology found in an increasing number of classrooms, and under controlled conditions, would be able to quickly and precisely measure identified students' level of content mastery and fluency on a variety of tasks.

Further research is needed to investigate the saliency of a wider range of chronometric tasks in determining learning disorders. However, given the significant time and related cost advantages provided by utilization of chronometric tasks, it appears they would be of great potential benefit to practitioners, either as effective and quick screening instruments or as a content-specific supplement to more comprehensive diagnostic batteries.

REFERENCES

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: Author.

Ceci, S. J. (1990). On the relation between microlevel processing efficiency and macrolevel measures of intelligence: Some arguments against current reductionism. *Intelligence, 14,* 141–150.

Cisero, C. A., Royer, J. M., Marchant, H. G., & Jackson, S. J. (1997). Can the Computer-based Academic Assessment System (CAAS) be used to diagnose reading disability in college students? *Journal of Educational Psychology, 89,* 599–620.

Deary, I. J. (2000). Simple information processing and intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 267–284). New York: Cambridge University.

Donders, F. C. (1969). Over de snelheid van psychische processen. Onderzoekingen gedann in het Psysiologish Laboratorium der Utrechtsche Hoogeschool. In W. G. Koster (Ed. & Trans.), *Attention and performance II* (Vol. 30, pp. 412–431). (Original work published 1868, 1869)

Fleischner, J. E., & Manheimer, M. A. (1997). Math interventions for students with learning disabilities: Myths and realities. *School Psychology Review, 26,* 397–413.

Galton, F. (1883). *Inquiries into human faculty and its development.* London: Macmillan.

Geary, D. C. (1994). *Children's mathematical development: Research and practical applications*. Washington, DC: American Psychological Association.

Gonzalez, J.E.J., & Espinel, A.I.G. (1999). Is IQ-achievement discrepancy relevant in the definition of arithmetic learning disabilities? *Learning Disability Quarterly, 22,* 291–301.

Hayes, F. B., Hynd, G. W., & Wisenbaker, J. (1986). Learning disabled and normal college students' performance on reaction time and speed classification tasks. *Journal of Educational Psychology, 78,* 39–43.

Jensen, A. R. (1985). Methodological and statistical techniques for the chronometric study of mental abilities. In C. R. Reynolds & V. L. Wilson (Eds.), *Methodological and statistical advances in the study of individual differences* (pp. 51–116). New York: Plenum.

Jensen, A. R. (1987). Mental chronometry in the study of learning disabilities. *The Mental Retardation & Learning Disability Bulletin, 15,* 67–88.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.

Kazdin, A. E. (2002). *Research design in clinical psychology* (4th ed.). Boston: Allyn & Bacon.

Knoop, A., Beaujean, A. A., & Holliday, G. (2005, March). *Descriptive differences in mathematics performance in the early lifespan: Kindergarten through college*. Presentation at the annual Lifespan Development Initiative, Columbia, Missouri.

Kranzler, J. H. (1994). Application of the techniques of mental chronometry to the study of learning disabilities. *Personality and Individual Differences, 16,* 853–859.

Kranzler, J. H., & Jensen, A. R. (1989). Inspection time and intelligence: A meta-analysis. *Intelligence, 13,* 329–347.

Kulak, A. G. (1993). Parallels between math and reading disability: Common issues and approaches. *Journal of Learning Disabilities, 26,* 666–673.

Luciano, M., Wright, M. J., Geffen, G. M., Geffen, L. B., Smith, G. A., & Martin, N. G. (2004). A genetic investigation of the covariation among inspection time, choice reaction time, and IQ subtest scores. *Behavior Genetics, 34,* 41–50.

Meyer, D. E., Osman, A. M., Irwin, D. E., & Yantis, S. (1988). Modern mental chronometry. *Biological Psychology, 26,* 3–67.

Neter, J., Kutner, M. H., Nachtsheim, C. J., & Wasserman, W.

(1996). *Applied linear statistical models* (4th ed.). Boston: McGraw-Hill.
Posner, M. I. (1978). *Chronometric explorations of mind.* Hillsdale, NJ: Lawrence Erlbaum.
Royer, J. (1999). *The Computer-based Academic Assessment System* (Windows version). Belcherton, MA: Educational Help.
Spearman, C. (1904). "General intelligence": Objectively defined and measured. *American Journal of Psychology, 15,* 201-292.
Swanson, H. L. (1987). Information processing theory and learning disabilities: A commentary and future perspective. *Journal of Learning Disabilities, 20,* 155–166.
Vernon, P. A. (Ed.). (1987). *Speed of information-processing and intelligence.* Norwood, NJ: Ablex.
Wechsler, D. (1997a). *Wechsler Adult Intelligence Scale-Third Edition.* San Antonio, TX: The Psychological Corporation.
Wechsler, D. (1997b). *Wechsler Memory Scale-Third Edition.* San Antonio, TX: The Psychological Corporation.
Whyte, J., Curry, C., & Hale, D. (1985). Inspection time and intelligence in dyslexic children. *Journal of Child Psychology and Psychiatry, 26,* 423–428.
Woodcock, R., McGrew, K., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement.* Itasca, IL: Riverside Publishing Company.

FOOTNOTES

1. A logit value is nonlinear transformation of the probability values (see Footnote 2). It is defined as follows:

$$logit_i = ln(\frac{\pi_i}{1-\pi_i})$$

where π is the *i*th individual's probability value and *ln* is the natural logarithm.

2. The probability, denoted by the Greek letter pi, π, that a given individual will be classified into Group 1 (case group), as opposed to Group 0 (clinical comparison control group), is defined as follows:

$$\pi_i = \frac{\exp(b_0 + \sum_{i=1}^{n} b_j X_{ij}}{1 + \exp(b_0 + \sum_{i=1}^{n} b_j X_{ij}}$$

where *i* is a specific individual, b_0 is the model's intercept and b_j are the coefficients for the scores on variable X_j (j=1, 2, . . .). It ranges from 0 to 1. For the model in this analysis, the b_0 value is -13.840 and the single variable coefficient for the CAAS RT is 1.530.

3. A more powerful study would have used a combination of tasks as well as a community comparison group (i.e., a group of college students with no math problems). Due to the constraints of the larger study, this was not possible at the time of data collection. Because of the results of the initial study, data are currently being gathered on the community comparison group as well as other math-based chronometric tasks.

AUTHOR NOTE

This material is based upon work supported by the National Science Foundation under Grant No. 0099216 (Principal Investigator: Gregory A. Holliday).

Correspondence concerning this article should be addressed to: A. Alexander Beaujean, Assessment and Consultation Clinic, 205 Lewis Hall, University of Missouri, Columbia, MO 65211; abeaujean@sigmaxi.org