

## ON THE TIME TO SUBTRACT: DONDERS REVISITED

William M. Petrusic

*Department of Psychology, Carleton University, Ottawa, Ontario, Canada,  
[bill\\_petrusic@carleton.ca](mailto:bill_petrusic@carleton.ca);*

*Extending Muensterberg's (1889) application of Donders' (1868) subtraction method, participants were required to subtract and compare the difference with a number (e.g., 9-7 vs. 3) and on other trials to simply compare comparable numbers (e.g., 2 vs. 3). The time to subtract and compare depended on the difference between the numbers compared. Problem size effects were also shown to depend on decisional difficulty, permitting the rejection of the intuitively compelling retrieve and compare model. The findings provide support for a model of the subtract and compare task in which numbers are represented as noisy extents, and the problem size effect is captured through the assumption that the variability of the representation is proportional to the square of its expectation. The simple discrete evidence accumulator is used to model the comparison process. These assumptions taken together permit a quantitative account of both the response time and the accuracy data.*

Early in the history of experimental psychology, Münsterberg (1889) extended Donders' (1860) complication-subtraction technique in order to obtain estimates of the time taken to activate knowledge stored in long-term memory. One domain he investigated, among several others, involved knowledge for simple arithmetic facts and he sought to determine the time taken to perform a simple subtraction operation. To this end, on some trials he required observers to subtract and then compare the difference with a number (e.g., "was ist weniger funfzehn oder zwanzig minus acht?"; "which is smaller fifteen or twenty minus 8?") and on other trials, observers simply compared comparable numbers (e.g., 15 vs. 12). Subtracting response times for the simple binary comparison task from those in the subtract and compare, compound task, putatively provided an estimate of the time to perform the subtraction operation.

Münsterberg obtained only very preliminary data and he did not present response times for the subtract and compare task explicitly and, thus, the success of his method could not be fully gauged. Hence, one of the purposes of the present experiment was to explore Münsterberg's ideas further by using both the compound and simple comparison tasks to determine the time to subtract and, more generally, to determine the utility of Münsterberg's approach to the study of mental arithmetic.

The present experiments provide an alternative means for the exploration of the fundamental properties of mental arithmetic, subtraction in particular. As will become evident the noisy-analogue-model provides a rather detailed characterization of both the processes involved in the compound subtract and compare task and the basic features of the subtraction process. One of our objectives is to determine if the properties of subtraction explicated in the compound subtract and compare task, coincide with those obtained with verification and verbal reports of solutions.

***Models for the compound --- subtract and compare task***

More generally, the experiment was conducted to explore the manner in which the decisional systems in comparative judgment and the knowledge activation systems involved in simple mental arithmetic interface. To this end, we consider two alternative theoretical possibilities, which we refer to as the additive stages model and the interactive evidence accrual model.

***The additive stages model.*** Of course, we should expect the “pure insertion” and the consequent subtraction of RTs, based on Donders’s view, to fail for a host of conceptual and methodological reasons that have been well documented throughout the history of psychophysics and RT. However, following on Sternberg’s additive factors methodology we might consider an intuitively plausible model, which posits separate stages of processing in the compound subtract and compare task. One version of this model, a variant and extension of the model first developed by Parkman (1972) for verification in multiplication, might posit the following stages of processing: i) compute or retrieve the difference in the subtraction problem and maintain it in memory; ii) encode the instruction and the comparison number and initiate the number comparison process; iv) execute the appropriate motor response.

The retrieve and compare model asserts that the retrieval operation must precede the comparison and, consequently, factors influencing the decision-comparison stage must not interact with those influencing the retrieval stage. Notably, the difference in RTs between the subtract and compare and the compare only conditions must not depend on the difficulty of the comparison. Furthermore, in the present experiment we also factorially vary the magnitude of the operands (problem size effect) and the difficulty of the comparison (the split effect). The retrieve and compare model is clear in not permitting an interaction of these two factors.

***An evidence accrual noisy analogue view.*** Restle (1970), in his landmark analyses of mental addition and comparison provided data in support of the idea that the representation of numbers might be likened to that of line lengths. The process of mental addition, as in the axiomatizations of extensive measurement (see Krantz, Luce, Suppes and Tversky, 1971), can be likened to the laying off of rods, (end to end concatenation). The process of adding and comparing, then amounts to simply comparing two extents, one obtained from the concatenation operation and the other arising from the number that is to be compared with the sum. These ideas are formalized in the case of mental subtraction below for the case where the participant is required to compare the difference between numbers,  $m$  (the minuend) and  $s$  (the subtrahend), with the number  $c$ ; for example, “which is larger, 9-7 or 3?” with the following assumptions.

*Assumption 1.* Numbers,  $X$ , are represented as random variables,  $\mathbf{X}$ , and in the present context, we assume they are Gaussian distributed with expectation,  $E(\mathbf{X})=\mu$  and variance,  $VAR(\mathbf{X})=\sigma^2$ .

*Assumption 2.* The subtraction operation is based on the random variable  $\mathbf{R}=\mathbf{M}-\mathbf{S}$ , where  $\mathbf{M}$  and  $\mathbf{S}$  are random variables representing minuend and subtrahend, respectively and the comparison number  $C$ , which is represented by the random variable,  $\mathbf{C}$ . The variability of the subtraction operation is given by  $VAR(\mathbf{R})=VAR(\mathbf{M}) + VAR(\mathbf{S})$ . The variability of the comparison process is given by  $VAR(\mathbf{R}-\mathbf{C})=VAR(\mathbf{C})+VAR(\mathbf{R})$ , where  $\mathbf{R}$  is a random variable representing the remainder,  $R$ .

*Assumption 3.* The overall variability in the subtract and compare task is the sum of the variability of the comparison task and the variability of the subtraction operation.; i.e.  $VAR(\mathbf{R})+VAR(\mathbf{C})+VAR(\mathbf{M})+VAR(\mathbf{S})$ .

*Assumption 4.* The problem size effect is captured by assuming that  $VAR(X)=\sigma^2 =k E^2(X)$ , where  $k>0$ . For convenience, we assume,  $E(\mathbf{X})=X$ .

*Assumption 5.* The comparison process is based on a Thurstonian (1927), unbiased, decision rule:  $R$  is greater than  $C$  if and only if  $\mathbf{R}-\mathbf{C}>0$ . Letting  $p$  denote *Probability*( $\mathbf{R}-\mathbf{C}>0$ ),  $p=F(z)$ ,

where  $\mathbf{Z}$  is a standardized Gaussian random variable and  $F$  denotes the distribution function. In particular:

$$Z=(R-C)/(k(M^2 + S^2 + R^2 + C^2))^{1/2}. \quad (1)$$

*Assumption 6.* The decision is based on an evidence accrual process. Consider the simplest evidence accrual model of comparative judgments – Pike’s (1965) discrete accumulator. Letting  $r_A$  and  $r_B$  denote the criterion (recruitment) levels for responses  $A$  and  $B$ , respectively, and  $p$  the probability of evidence favouring the response  $A$ , the probability of response  $A$  is given by the expression,

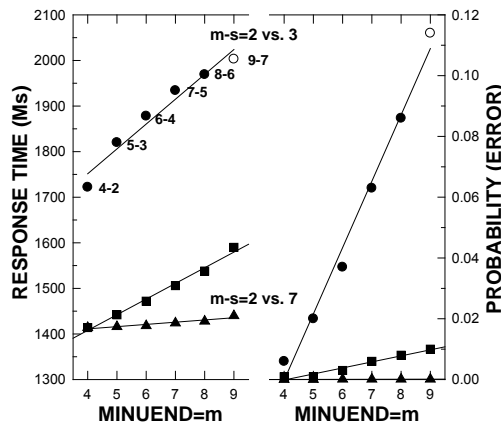
$$P_A = \sum_{j=1}^{r_A-1} (r_A + j - 1)! p^{r_A} q^j / j!(r_A - 1)!, \quad (2)$$

where the probability  $p$  in the above expression is given by  $F(z)$ ,  $z$  is defined by Equation 1, and  $F$  is the cumulative Gaussian distribution function. The mean number of accruals conditional on the occurrence of response  $A$ , is given by the expression,

$$E(N | R_A) = \frac{1}{P(R_A)} \sum_{j=1}^{r_A-1} (r_A + j - 1)! p^{r_A} q^j / (j - 1)!(r_A - 1)!, \quad (3)$$

and it is assumed that  $RT(R_A)=aE(N/R_A)+b$ . The constant  $a$  reflects the duration of each accrual and  $b$  non-decisional, residual components.

In summary, these assumptions, taken together, predict that the split effect will be greatly enhanced in the subtract and compare condition relative to the simple compare condition. Further, the noisy-analogue evidence accrual view predicts that the efficacy of arithmetic processing will depend on the difficulty of the numerical comparison required. In particular, problem size effects should be accentuated as the difficulty of the comparison increases (i.e., as split decreases). These predictions are illustrated in Figure 1 for both the probability of correct and the average correct RT measures.



*Figure 1.* Mean correct response times (left panels) and Probability (Error) (right panels) to subtract and compare, with minuend and subtrahend varying and minuend minus subtrahend fixed at 2 for splits of 1, 3 and 5 for the top, middle, and bottom panels, respectively, as predicted by the noisy analogue theory. These theoretical predictions are based on setting  $r_A = r_B = 7$ ,  $k = 0.04$  and  $a = 200$  ms.

### Method

*Subjects.* Twenty-six male and twenty-two female first year psychology students participated for course credit for a single one hour session.

*Apparatus.* Stimuli and instructions were presented on an Amdek-310A video monitor that was controlled by an IBM-PC/XT compatible computer. The computer also controlled trial presentation, event sequencing, randomization, and the recording of responses and response times

*Design and procedure.* The experiment was comprised of 196 trials of the *compound*, subtract and compare form (e.g., “which is larger: 9-7 or 3?”) and 160 *simple* binary comparisons (e.g., “which is larger: 2 or 3?”). , there were 49 compound pairs in the experimental set and a further 2 pairs were used for practice trials. Each of the 49 compound pairs was used in each of two presentation orders (left vs. right), and the two instructions (“choose the larger” and “choose the smaller”). This factorial combination of pairs, instructions and orders thus resulted in the 196 compound trials.

The 196 compound trials were also comprised of three levels of *split* (the difference between the numbers compared): 1 (e.g., 4-2 vs. 3), 3 (e.g., 9-2 vs. 4) and 5 (e.g., 8-1 vs. 2) and each level of split was used with each of the two instructions and the two presentation orders. As well, for each compound comparison split, comparable simple comparisons were used (e.g., 4-2 vs. 3 and 2 vs. 3).

Each trial was initiated with the presentation of an instruction (LARGER/SMALLER), centered and near the top of the screen. One second after the presentation of the instruction the digits for the comparison trial were presented, centered horizontally and vertically on the screen. They and the instruction remained on the screen until the subject responded. The screen was then cleared and the next trial followed after a 2 s inter-trial interval. Instructions emphasized both speed and accuracy.

## Results

The data from two participants who failed to use the correct buttons on the “Mouse” were eliminated. Trials with response times longer than 5 s and less than 200 ms were discarded. These trials accounted for 0.4 % of the simple comparison trials and 1.2 % of the compound, subtract and compare, trials. Except for explicit examinations of error RTs, mean RT for correct responses for each cell of the design for each subject served as the dependent variable and the probability of an error served as the dependent variable in analyses of accuracy. The findings are presented in two main sections. The first examines the main features of the comparison process and the second explores the properties of subtraction. Throughout, level of significance was set at 0.05 and the Huynh-Feldt epsilon adjusted degrees of freedom were employed although the degrees of freedom and mean square errors provided are those based on the design.

### *Global analyses*

*Response times.* On average, RT s for simple comparisons were 885.66 ms and 1561.59 for compound comparisons and the requisite subtraction operation reliably increased comparison time ( $F(1, 45)=387.52$ ,  $MS(Error)=81351.77$ ). As well, RT s increased significantly as the difference between the numbers compared decreased ( $F(2, 90)=75.72$ ,  $MS(Error)=13086.66$ ).

Contrary to Münsterberg's idea, the time to subtract could not be determined by subtracting the reaction times in the comparison alone condition from the times to subtract and compare. The interaction between type of comparison and split was statistically reliable ( $F(2, 90)=57.12$ ). As is evident in Figure 2 distance effects in the subtract and compare condition were greatly enhanced relative to comparison alone, thereby, establishing an interdependence among memory activation processes in subtraction and decisional processing in comparison as predicted from the evidence accrual view. It should also be noted that error responses were faster than correct responses for both compound and simple comparisons, suggesting subjects were willing to sacrifice accuracy for speed in both tasks.

*Error rates.* On average, errors occurred on 0.023 and 0.053 of the trials for simple and compound comparisons, respectively, and the requirement to also subtract reliably increased error rates ( $F(1, 45)=33.51$ ). Paralleling the main effect of RT for split, error rates reliably increased as split decreased ( $F(2, 90)=32.44$ ). Also paralleling the RT analyses, the interaction between split and type of comparison ( $F(2, 90)=34.74$ ) was reliable. Thus, the

strict additivity of the subtraction operation and the decisional comparison processing is denied on the basis of both the error and the RT analyses, contrary to the simple stage, retrieve and compare model.

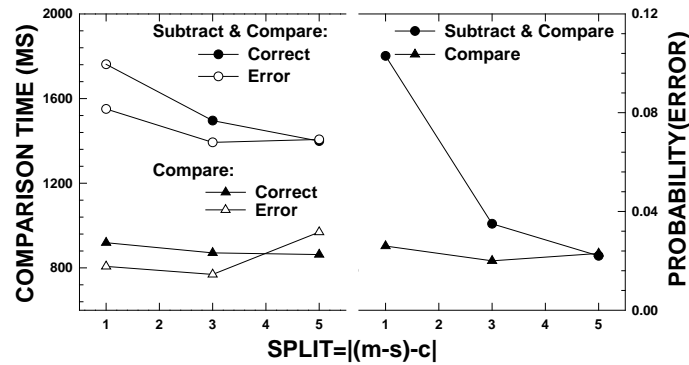


Figure 2. Mean correct and error times for the subtract and compare and the compare only tasks as a function of split and mean probability of an error for the subtract and compare and the compare only conditions.

*Properties of the subtraction process*

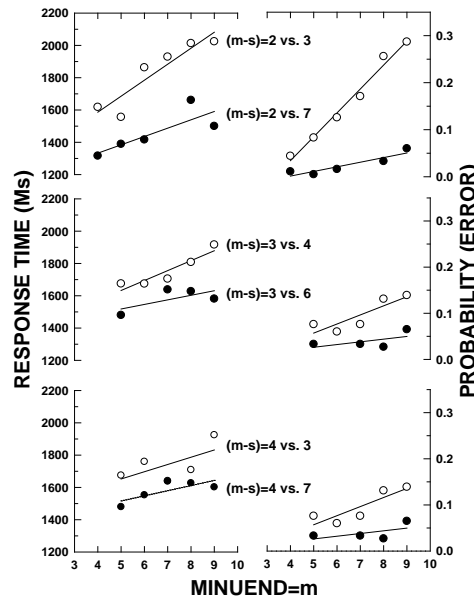


Figure 3. Mean correct response times to subtract and compare and Probability (Error) with minuend and subtrahend varying and minuend minus subtrahend fixed at 2 for splits of 1, 3 and 5.

*Magnitude effects in subtraction.* The plots in Figure 3 also show that the times to subtract and compare, with comparative decision difficulty held constant, reveal a clear dependence on the magnitude of the numbers subtracted. Letting  $d$  denote the difference,  $m$ , the minuend, and  $s$ , the subtrahend, then when  $d=m-s=2$  is compared with 3,  $m=9$  and  $s=7$  (i.e., 9-7 vs. 3), response times are maximal and errors occur on 29.9% of the trials, precisely as predicted by the noisy analogue model (see Figure 1). As both the minuend and the subtrahend decrease (e.g., to  $m=4$ ,  $s=2$ ; with 4-2 vs. 3), response times decrease (approximately linearly) as do error rates. Similar effects are also evident when decision difficulty is held constant in comparisons of  $m-s=3$  vs. 4 (middle panel) and with  $m-s=4$  vs. 3 (bottom panel). Moreover, these problem size effects are evident with both the more difficult comparisons with a split of one and with the easier comparisons with a larger split of five. Thus, the dependence of the time to subtract on the magnitude of the numbers involved is

rather robust and readily evident with compound comparisons with the difficulty of the comparison held constant.

*Dependence of the problem size effect on the difficulty of the comparison.* The plots in Figure 3 also show that the slopes of the problem size effect, for both RTs and errors, are substantially reduced when larger splits are involved; i.e., when the comparison is easier. Thus, the problem size effect is clearly dependent on the difficulty of the comparison, precisely as predicted by the noisy analogue evidence accrual model and contrary to the retrieve and compare model. For the plots shown in the upper panels in Figure 2, with (m-s)=2 vs. 3, an RT based ANOVA, employing orthogonal polynomials, resulted in a significant overall linear trend ( $F(1, 44)=32.51$ ), affirming problem size effects for subtraction. Importantly, the interaction between the linear trend component and split was reliable ( $F(1, 44)=5.03$ ), reflecting the dependence of the problem size effect on the difficulty of the comparison. Comparable, statistically reliable effects occurred with the error rate measure. Moreover, ANOVAs with both the RT and error rate measure for the problem size effects examined in the middle and bottom panels in Figure 1, resulted in each case (except for the accuracy measure in the bottom panel) in statistically reliable linear trends and interactions of this linear trend with the split factor.

### ***Summary and conclusions***

The evidence is decisive against the view that the time to subtract and compare is equal to the time to subtract plus the time to compare, contrary to Münsterberg's (1892) surmise and the intuitively compelling retrieve and compare model, popular in verification tasks. Rather, taken together with the enhanced split and semantic congruity effects and, especially the dependence of the problem size effect on split, considerable support is provided for the evidence accrual model with noisy memory based inputs. The present findings of problem size effects in the subtract and compare task also converge nicely with those obtained by Restle (1970) in his add and compare task and provide further support for his view that numbers are represented as noisy linear magnitudes. In fact, the compound and compare task serves to magnify problem size effects, precisely as predicted by the noisy analogue evidence accrual model, where numbers are indeed represented as noisy linear magnitudes.

### ***References***

- Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement* (Vol. 1). New York: Academic Press.
- Restle, F. (1970). Speed of adding and comparing numbers. *Journal of Experimental Psychology*, **83**, 274-278.
- Münsterberg, H. (1889). *Beitrage zur experimentellen psychologie*. Heft 1. Freiberg: von J.C.B. Mohr. pp. 94-106.
- Parkman, J.M. (1972). Temporal aspects of simple multiplication and comparison. *Journal of Experimental Psychology*, **95**, 437-444.

### ***Acknowledgement***

This work was supported by a Natural Sciences and Engineering Research Council of Canada Discovery grant to William M. Petrusic