EFFECT OF SPEED AND ACCURACY INSTRUCTIONS ON CONFIDENCE JUDGMENTS FOR MAGNITUDE ESTIMATES*.

Ana Garriga-Trillo	C. Díaz-Berciano	I. Serrano
UNED	CES San Pablo CEU	UCM
Madrid, Spain	Madrid, Spain	Madrid, Spain

Abstract

Accuracy and speed instructions have been recently and successfully used to determine the locus of confidence judgments for detection and discrimination tasks (Baranski & Petrusic, 1998; Petrusic & Baranski, 2000). No such research has been done using magnitude estimation. This paper studies, using an absolute magnitude estimation task and four experimental conditions, the effects of speed and accuracy instructions on responses, proportion of correct responses, decisional response times, confidence judgments and confidence response times. The proportion of correct responses used has been previously defined in Garriga-Trillo (1996). Results show greater effects for the accuracy condition than for the speed one in all the five dependent variables considered. It seems that subjects not only apply both instructions to estimate magnitudes but also to estimate confidence. When confidence is demanded, either under accuracy or speed conditions, mean decisional response times are significantly longer than for the non-confidence condition. Confidence judgments could be, to a certain extent, processed with the response.

Magnitude estimation involves a detection and a discrimination phase, previous to giving a numerical response to quantify stimulus magnitude. Subjects first have to detect the stimulus, then discriminate between stimuli and, finally assign a number to the perceived stimulus magnitude. Following this inclusion mechanism and although recognizing a more complex undertaking, studying aspects from detection and discrimination tasks can shed light on magnitude estimations. This approach has been taken by Garriga-Trillo, Villarino, González-Labra & Arnau (1993, 1994) to begin studying confidence judgments of magnitude estimates concerning the underconfidence-overconfidence effect and considering studies with detection and discrimination tasks. Díaz & Serrano (2000) and Serrano & Díaz (2000) have considered issues from previous research done with discrimination tasks to study the same effect with a magnitude estimation task.

Very recently, according to Petrusic and Baranski (2000), theories trying to explain the basis for confidence judgments, within a general framework of detection and discrimination, have been developed. Balakrishnam & MacDonald (2000) mention that confidence judgments may be the result of an information accumulation process or a derivative of such a process produced after the discrimination decision is made. Although most models present the locus of confidence after the decisional process is attained (its locus is then post-decisional) researchers like Baranski & Petrusic (1998) present evidence of a pre-decisional locus under some conditions and a post-decisional locus under other conditions. These conditions were mainly related to accuracy and speed instructions. Although their work, and that of others, concern detection and discrimination tasks, as we stated at the beginning, it is useful to implement their ideas to continue studying confidence judgments of magnitude estimates. This work pretends to study the effects of accuracy and speed instructions on confidence judgments of magnitude estimates using a sensory task, the estimation of rectangles areas, having in mind a future search for a possible locus of confidence judgments of magnitude estimates.

METHOD

Subjects. Sixty-four volunteer subjects (14 males and 50 females; mean age = 23.4 years; age range = 18-55 years) participated in the experiment. All the subjects had normal or corrected-to-normal vision and were naive to the nature and aims of the experiment. All participants were tested individually.

Apparatus. The stimuli were presented on a 14-inch screen monitor, which had a resolution capability of 800 pixels horizontally by 600 pixels vertically. A Pentium II PC was used and a specific programmed software (Serrano, 1999) controlled instructions, event sequencing, randomisation, and the recording of responses, response times, confidence ratings and confidence response times. The computer keyboard was used as an input device.

Stimuli. Eight rectangles of different areas were displayed centred on the computer screen. They appeared in white on a black background and were generated making their length 1 cm longer than their width both within the rectangles and between them. For the experiment, the areas considered were 2, 6, 12, 20, 30, 42, 56 and 72 squared centimetres. Another four stimuli were used for the practice trial session, but using a 2 centimetre difference between length and width. Their areas were 3, 15, 35 and 63 squared centimetres. All stimuli were randomised within blocks. The display was presented at eye level and at a viewing distance of approximately 50 cm. Previous to the rectangles' display, a white circle appeared on the centre of the screen to focus the subjects' attention.

Procedure. Four experimental conditions were defined considering four different sets of instructions (accuracy-confidence, speed-confidence, accuracy-no confidence, speed-no confidence). Subjects were randomly assigned to one of the four groups. The instructions were presented on the screen. All groups had to estimate the rectangles´ areas using the absolute magnitude estimation technique both in the trial session and in the three experimental blocks. The speed instructions encouraged observers to respond immediately after the rectangles disappeared from the screen. In the accuracy instructions observers were instructed to be accurate in their estimations of the rectangles´ areas. The confidence groups were required to give a confidence rating, from 1 to 100, for their magnitude responses. A confidence value of 100 indicated absolute certainty, and 1 indicated no certainty at all. When the participant had read and understood the instructions, s/he pressed the "Enter" key and the experiment began. The white circle appeared on the centre of the screen to focus the subjects´ attention. Then, the first rectangle appeared centred on the screen. When the subject was ready to estimate its area s/he pressed the "Enter" key and the primary decision time was recorded. This primary decision time is then the time measured from stimulus onset to pressing the "Enter" key. The screen was cleared and a question asking for the area of the rectangle then appeared on the screen. Observers indicated their estimated area by moving the index finger of their preferred hand from the "Enter" key to the number response keys. This secondary decision time (the time measured from starting with the number keys to ending giving the numerical answer) and the number given were also recorded. The program also recorded the sum of both primary and secondary decision times, the decisional response time (Decisional RT). If a confidence rating was demanded from the subject, observers pressed the "Enter" key and moved the same finger from the "Enter" key to the numerical keys to express their confidence by pressing their response using the same numerical keys. Thus the time from the depression of the “Enter” key to the depression of the last number from the numerical keys was the time to determine confidence, or confidence response time (Confidence RT). The confidence judgment and the confidence RT were also recorded. All RTs were measured in seconds.

RESULTS and DISCUSSION

Our data has been analyzed both descriptively and inferentially. Figures 1-5 show plots in which accuracy and speed are one of the independent variables considered. The other independent variable is either stimulus magnitude or blocks. A new independent variable is considered in Figure 6: confidence and no confidence conditions. The dependent variables were response, proportion (or percentage) of correct responses, decisional RT, confidence ratings and confidence RT. The proportion of correct responses (or its percentage) was calculated as defined in Garriga-Trillo (1996) using an index similar to Kendall´s Tau.

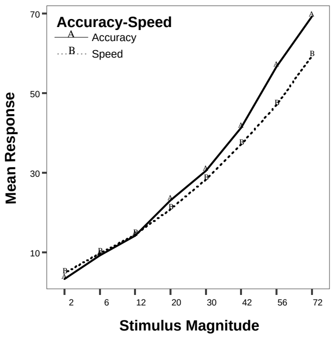

Figure 1. Mean response for each stimulus magnitude under accuracy-speed conditions.

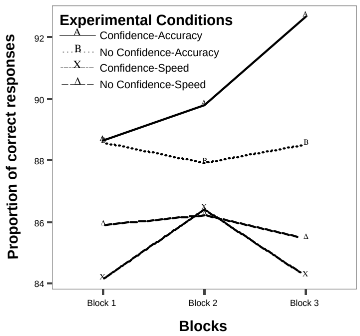

Figure 2. Proportion of correct responses for each block under four experimental conditions.

In Figure 1 we can observe an ascending trend in responses as stimulus magnitude increases. There are significant differences between responses, $F(7,1528)=134$, $p<0.0001$, Power=1. Subjects discriminate between stimuli. Also mean responses (each point represents 192 judgments) are significantly greater in the accuracy condition than in the speed one, $F(1,1534)=6.55$, $p<0.01$, Power=0.73. The interaction between the independent variables was not significant.

Figure 2 considers four conditions and three blocks as independent variables. The differences in proportion of correct responses between blocks is not significant, $F(2,189)=0.19$, $p>0.83$, Power=0.08. No learning effect is detected. The differences in percentage of correct responses between the four conditions is significant, $F(3, 188)=3.6$, $p<0.01$, Power=0.79, although all the percentages are high. This may mean that the task was easy. The highest percentages are obtained in confidence-accuracy (mean=90.4) and the smallest ones in confidence-speed (mean=84.9). The mean difference, using the Tamhane statistic, is significant for $p<0.03$. The other differences between conditions are not significant. The interaction between the two independent variables was not significant.

In Figure 3 one can observe an ascending trend in mean decisional RT as stimuli increases. It takes longer to asses a large area than a small one. One could say that it is more difficult to assess a large area than a smaller one. In fact, significant differences in decisional response time between stimuli are found, $F(7, 1528)=8.4$, $p<0.0001$, Power=1. Mean decisional RT values go from 6.95-10.56 seconds. There are also significant differences in mean decisional response time between the accuracy and speed conditions, $F(1,1534)=69$, $p<0.0001$, Power=1. Mean decisional RTs for accuracy are longer than for speed. In both conditions mean decision RT increases as stimulus magnitude increases. The interaction between the independent variables is not significant.

In Figure 4, there is a descending tendency, seen for both accuracy and speed, in confidence ratings as stimulus magnitude increases. We have just seen the opposite trend with decisional RT and stimulus magnitude. Combining both results with an example within the accuracy condition, the largest area (72 cm^2) has the longest mean decisional RT and the lowest confidence rating. Baranski & Petrusic (1998) found a similar result with a discrimination task. The accuracy condition has the highest mean confidence ratings compared with the speed condition. This difference between conditions considering mean confidence ratings is significant, $F(1, 766)=25$, $p<0.0001$, Power=0.999. Mean confidence ratings for the different stimuli are different, $F(7, 760)=2.9$, $p<0.005$, Power=0.93. The interaction between accuracy-speed and stimulus magnitude is not significant.

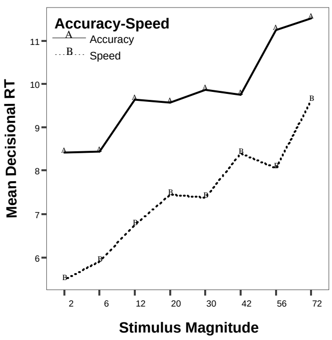

Figure 3. Mean decisional response times for each stimulus magnitude under accuracy-speed conditions.

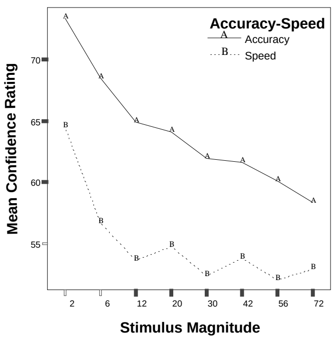

Figure 4. Mean confidence ratings for each stimulus magnitude under accuracy-speed conditions.

Figure 5 represents an uneven path for stimulus magnitudes and confidence response times in both accuracy and speed conditions. Nevertheless, there are significant differences in accuracy-speed conditions considering confidence RT, $F(1,766)=6.66$, $p<0.01$, Power=0.73. The accuracy condition has the longest confidence RT. No significant differences can be found between stimulus values regarding mean confidence RT, $F(7, 760)=0.4$, $p>0.9$, Power=0.18. No significant interaction effects can be found.

Figure 6 presents stimulus magnitude and confidence-no confidence instructions as independent variables. Mean decisional response time is the dependent variable. As in Figure 3 for accuracy and speed, and being the other independent and the dependent variables the same, the same significant differences in mean decisional RT for the different stimuli are found. There are also significant differences in mean decisional RT between confidence conditions, F(1, 1534)=5.3, p<0.02, Power=0.64. The confidence condition presents longer decisional RT (the difference between groups is not as large as it was for the accuracy-speed conditions). This could imply that confidence judgments began to be processed before the subject's decision of assigning a specific number to the perceived magnitude of the stimulus. Comparing the ranges of mean decisional RT and mean confidence RT we find that the first one goes from 6.95-10.56 and for mean confidence RT it goes from 3.05-3.98. Since both tasks involve giving a number to a perceived magnitude there should not be such a difference in their response times. A plausible explanation could be that decisional RT includes part of the process of giving a confidence rating. Calculating Pearson's Product Moment Correlation between overall mean confidence RT and mean decisional RT, one obtains 0.7, which is significant for $p<0.0001$. Almost the same value ($r_{AC} = 0.72$, $p<0.002$ and $r_{SP}=0.71$, $p<0.002$) is obtained within the accuracy (AC) and speed (SP) conditions.

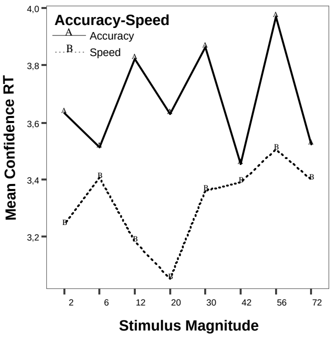

Figure 5. Mean confidence response times for each stimulus magnitude under accuracy-speed conditions.

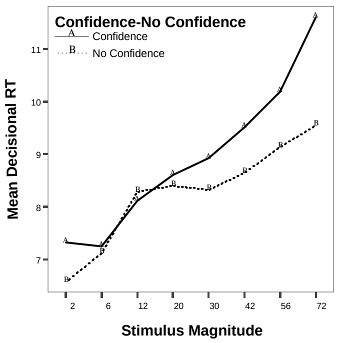

Figure 6. Mean decisional response times for each stimulus magnitude under confidence-no confidence conditions.

As a conclusion of our analyses one can state that accuracy is a very relevant variable when working with confidence. Interpreting the r-value between confidence RT and decisional RT, 49% of the variance of decisional RT is explained by the variance of confidence RT. It seems that decisional RT includes some time to determine confidence, as Baranski & Petrusic (1998) suggest when considering a discrimination task. Theories assuming the unique existence of a post-decisional locus for confidence judgments can be challenged, as stated by Petrusic & Baranski (2000).

* This work was done as a doctoral joint research project. UNED, Madrid.

References

Balakrishnan, J.D. & MacDonald, J.A. (2000). Recalibrating subjective confidence ratings: Implications for scaling and measurement. In C. Bonnet (Ed.) *Fechner Day 2000* (pp. 85-90). Strasbourg, France: Université Louis Pasteur-International Society for Psychophysics.

Baranski, J.V. & Petrusic, W.M. (1998). Probing the locus of confidence judgments: Experiments on the time to determine confidence. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 3, 929-945.

Díaz, C. & Serrano, I. (2000). Juicios de confianza sobre estímulos geométricos: Aspectos temporales. En *Libro de Resúmenes* (pág. 36), Tercera Semana de Investigación. Madrid: Facultad de Psicología.

Garriga-Trillo, A. (1996). How could you say that magnitude estimates are correct? An application to calibration studies. In S. C. Masin (Ed.) *Fechner Day 96* (pp. 257-262). Padua, Italy: University of Padua-International Society for Psychophysics.

Garriga-Trillo, A., Villarino, A., González-Labra, M.J. & Arnau, M.A. (1993). Calibration of psychophysical judgments: Magnitude estimates. In A.Garriga-Trillo (Chair), *New Approaches to psychophysics - 2*. Session conducted at the Ninth Meeting of the International Society for Psychophysics, Palma de Mallorca, Spain.

Garriga-Trillo, A., Villarino, A., González-Labra, M.J. & Arnau, M.A. (1994). La calibración de juicios psicofísicos: Estimación de magnitudes. *Psicothema*, **6**, 3, pp. 525-532.

Petrusic, W.M. & Baranski, J.V. (2000). Effects of expressing confidence on decision processing: Implications for theories of RT and confidence. In C. Bonnet (Ed.) *Fechner Day 2000* (pp. 103-108). Strasbourg, France: Université Louis Pasteur-International Society for Psychophysics.

Serrano, I. (1999). Programa de realización de un experimento psicofísico con tiempos de reacción (Computer software). Programa no comercializado. Madrid: UNED.

Serrano, I. & Díaz, C. (2000). Juicios de confianza sobre estímulos geométricos: Aspectos de exactitud. En *Libro de Resúmenes* (pág. 36), Tercera Semana de Investigación. Madrid: Facultad de Psicología.