CONFIDENCE JUDGMENTS IN VISUAL TEMPORAL DISCRIMINATION: CROSS-CULTURAL STUDY

Irina Skotnikova Institute of Psychology Russian Academy of Sciences, Moscow, Russia

Thomas Rammsayer and Susanne Brandler University of Goettingen, Goettingen, Germany

ABSTRACT

An experimental study on confidence judgements in visual duration discrimination was conducted. Experimental design initially used in the Russian study was replicated for a German sample of observers. A comparison of performance between German and Russian samples of subjects was made. Overconfidence was found in German participants as well as in Russian ones. At the same time it's value was lower than in Russian persons and either close or higher as compared to that which was described in Canadian subjects, in distinction of underconfidence known in Swedish participants for the same discrimination level. Data discussed point to cross-cultural differences in confidence judgements in sensory and perceptual tasks.

A crucial issue discussed in extensive literature on "realism of confidence" is whether humans generally underestimate or overestimate their confidence in the correctness of their judgments (i.e., show underconfidence - UC or overconfidence - OC) as compared to actual performance. Some authors have found UC in sensory discrimination (Bjorkman et al., 1993, Garriga-Trillo et al., 1994; Olsson & Winman, 1996; Juslin & Olsson, 1997). They consider it to be a typical and universal bias in sensory tasks in distinction of general knowledge tasks where OC and Hard-Easy Effect (HEE - UC in easy tasks while OC in difficult ones) are known. Others have obtained OC at difficult discrimination levels (for PC (proportion of correct responses) <.8), and HEE as well (Baranski & Petrusic, 1994, 1995, 1999; Ferrel, 1995; Petrusic & Baranski, 1998; Stankov, 1998; Skotnikova, 1994: OC only while HEE was not studied). Data comparison showed that Swedish participants practically never made "certain" judgments (Olsson & Winman, 1996; Juslin & Ollson, 1997) while Canadian and USA subjects made them between 20% and 40% (Baranski & Petrusic, 1999) and Russian observers - in 90% cases (Skotnikova, 1994). Besides 46.3% of USA drivers judged themselves in the top 20% of drivers, whereas only 15.5% of Swedish drivers did the same (Svenson, 1981 - see: Baranski & Petrusic, 1999). Cultural and national differences in probabilistic judgments were observed (Whitkomb et al., 1995; Yates et al., 1997). On this base Baranski & Petrusic (1999) have recently suggested that confidence judgments in sensory and perceptual tasks may also display cultural and national differences. In the present study this problem was examined regarding to German persons in comparison with Russian ones.

METHODS

Visual discrimination between filled temporal intervals presented in pairs was studied. Experimental design initially used in the Russian study (Skotnikova, 1994; 2000) was replicated for a German sample of observers.

Stimuli. Stimuli were successive light flashes. One stimulus in each pair was 600 ms, another one was either the same or equal to (600 ms - Δt) where Δt -values were chosen individually in a preliminary session. Pairs of equal stimuli (600 ms each) and different ones (600 ms and (600 ms - Δt)) and a position of a longer stimulus in a pair were equiprobable and counterbalanced.

Procedure. Each observer estimated durations in each pair as "same" or "different" and then marked this response as "confident" or "unconfident" with the help of a keyboard. Instruction stressed accuracy but not speed of responses. Observation time was not limited. The experiment included 3 sessions.

1. Preliminary (p) session. A kind of a staircase up-and-down procedure was used. Δ t-values in region of 300÷100 ms were presented by decreasing and increasing steps (100÷2ms). The precision of Δ t-values presentation was 1ms. Individual Δ t-values corresponded to 70÷80% of correct responses (PC =.7 ÷ .8), i.e closed to just notiseable difference were determined. Each subsession (for each Δ t-value) consisted of 10 trials.

2. Training (t) session: 50 trials using a Δ t-value found in a preliminary session.

3. Main (m) session: 100 or (100 x 2) trials using a Δ t-value found in a preliminary session.

Participants. 15 voluntarees aged 22÷51 having normal or corrected to normal vision took part in the study: 8 males and 7 females; 7 students and 7 scientists of Goettingen University and 1 pensioner (former librarist). All 15 SS participated in the preliminary and training sessions, 6 of them - in 1 main session (m₁: 100 trials), 8 - in 2 ones (m₁ + m₂, i.e.: 2 x 100 trials), and 1 S - in a preliminary and training sessions only (his data obtained in a training session (t) of 50 trials were considered as data of a main session (m₁) of 100 trials).

Data processing. The following indices were calculated in each session: proportions of correct responses: PC; mean weighed confidence category used: $\text{Con} = 1/\text{N} \times (.5 \times n_1 + 1 \times n_2)$, where n_1 was a number of "unconfident" responses, n_2 - a number of "confident" responses, N- a total number of responses; bias (B) = Con - PC (see Ronis &Yates (1987) for the description of the indices). Data obtained were avereged across the group of observers by 2 ways: a) using data of sessions m_1 (i.e. across 100 trials x 15 SS = across 1500 trials); b) using data of sessions $m_1 + m_2$ (i.e. across 100 trials x 7 SS + (100 + 100 trials) x 8 SS = across 2300 trials). Results obtained were compared with data received in a sample of 29 Russian subjects. Statistical significance of differences between German and Russian data were estimated using Mann-Witney test on the base of "Statistica"-software. Stimuli presentation, responses recording and data processing were controlled by PC-AMD-K6-200 computer.

RESULTS

Results obtained are presented in the tables 1, 2. The following evidence was revealed. 1. Values of percent of correct responses (PC) were chosen individually in the range of .7÷.8 in the both groups and therefore the average values were found to be statistically the same: $MPC_{Ger} = .741, .742, \sigma = .034, .035; MPC_{Rus} = .731, \sigma = .050.$

Table 1. Data averaged across the German group and analogous data averaged across the sample of Russian subjects (Skotnikova, 1994, 2000)

Index	German		Russian
	Session m ₁	Sessions m1+m2	
MΔt,σ	168; 38	170; 39	143; 28
MPC, σ	.741; .035	.742; .034	.731; .050
Mcon, σ	.873; .010	.866; .0105	.950; .060
MB, σ	.123; .012	.118; .011	.219; .066

2. In the German group individual scores of the .7÷.8 -difference threshold (Δt) for the durations used were found in the range of 102÷275 ms, M Δt_{Ger} = 168, 170 ms,

 σ = 38, 39 ms. The corresponding Russian data were: Δt =100÷225 ms,

 $M\Delta t_{Rus} = 143 \text{ ms}, \sigma = 28 \text{ ms}.$ Values of Δt found in the German group were slightly (and at the same time significantly) greater than those in the Russian group.

Table 2. Comparison between German and Russian data.

Index	Session m ₁ , p<	Sessions m ₁ +m _{2,} p<
MΔt	.030	.006
MPC	.630 - insignificant	.429 - insignificant
Mcon	.0038	.000566
MB	.0012	.000068

3. In the German group individual values of mean confidence category used (Con) were found in the range of .650÷1.000. The same range of Con-values was discovered in the Russian group: .651÷1.000. At the same time the mean values of Con (MCon) were clearly lower in the German group than in the Russian one:

 $MCon_{Ger} = .866, .873 < MCon_{Rus} = .950; p < .0038, .000566.$

4. In the German group individual indices of confidence bias (B) were found in the range of $.120 \div .270$ while in the Russian group they were $.025 \div .341$. The average B-scores were

almost twice lower in the German group than in the Russian one: $MB_{Ger} = .118$, $.123 < MB_{Rus} = .219$; p < .001225, .000068. Only 1 German subject had shown negative B-scores, i.e. UC while the other 14 subjects - positive B-scores, i.e. OC.

DISCUSSION

At discrimination level corresponded to PC=.7-.8 OC was found in German persons which was either close or greater as compared to OC described in Canadian participants (Baranski & Petrusic, 1994, 1995, 1999; Petrusic & Baranski, 1998) but

as a rule lower than OC obtained in Russian subjects (Skotnikova, 1994, 2000). UC was found to be untypical for German persons in distinction of Swedish ones (Bjorkman et al., 1993; Olsson & Winman, 1996; Juslin & Olsson, 1997).

A review of data on confidence in different discrimination tasks has shown the following (see Skotnikova, 2000). A sign of confidence bias found in Swedish and in Canadian persons did not depend on a kind of stimuli used. For PC=.7+.8 and for 2,4 and more confidence categories used values of UC obtained in Swedish works were

-.145÷-.013 (Obrink, 1948; Bjorkman & Qvarsell, 1963; (Bjorkman et al., 1993; Olsson & Winman, 1996) for visual depth, speed, length of rectangles; weights; and values of OC obtained in Canadian studies were .010÷.072 for visual location and line length (Baranski & Petrusic, 1994, 1995, 1999; Petrusic & Baranski, 1998). Moreover, UC was again obtained in Swedish participants when the Canadian experiment on visual location and confidence estimation expressed in percent was replicated in Sweden (Olsson & Winman, 1996) in distinction of OC found in the initial Canadian study (Baranski & Petrusic, 1994). It points that confidence bias in sensory discrimination may depend rather on cultural-national differences, in according with Baranski & Petrusic (1999) suggestion, than on kinds of stimuli used.

At the same time a comparison between the German data from the one hand and the Swedish and Canadian data from the other hand looks to be not complete because of distinctions between psychophysical methods used. In the present study participants worked on the base of "Same-Different Method" in which they just dichotomized stimuli as same or different, like as in the initial Russian study. In distinction Swedish and Canadian subjects worked on the base of "Method of Constant Stimuli" and "Two Alternatives Forced Choice Method" in which they made an additional sensory operation as well: discriminated between the stimuli to be greater or lesser. On the base of comparison between the both procedures it was supposed that differentiated and therefore rather sensitive, precise character of "Greater-Lesser" discrimination may lead to a lesser overconfident bias, as compared to more rough "Same-Different"-discrimination (Skotnikova, 2000). It corresponds to an opinion that a decision of confidence is not merely a second "postdecisional" operation, which is an external one regarding to a first sensory decision but is included in it (when SS are "accuracy" but not "speed" instructed: Petrusic & Baranski, 1998). A significant role of a kind of a psychophysical procedure applied (and of a base duration as well) was revealed in temporal discrimination of auditory intervals. Long-term effects of practice were found for an adaptive procedure and brief base durations while no effects - for long durations using in this procedure and for temporal bisection paradigm (Brandler & Rammsayer, 1999). Therefore it looks to be important to take in account a specificity of a procedure used and a value of a base duration in psychophysical studies of confidence in time perception.

Besides, a negative time-order-error (TOE) was obtained in discrimination of 1000 msdurations which are rather close to those we used (Rammsayer & Wittkowsky, 1990; Hellstrom & Rammsayer, 2000). It has to be taken in account when we try to reveal a souce of the overconfident bias in German and Russian persons. It was shown that confidence judgments depend on stimuli presentation order, in particular on a space error (Baranski & Petrusic, 1999). If there is TOE in German and Russian data then it may induce a higher OC than in Canadian ones (Petrusic 2000 - personal communication). A further examination of the factors mentioned is under way.

REFERENCES

- Baranski, J.V. & Petrusic, W.M. (1994). The calibration and resolution of confidence in perceptual judgments. *Perception and Psychophysics*, 55, 412-428.
- Baranski, J.V. & Petrusic, W.M. (1995). On the calibration of knowledge and perception. Canadian Journal of Experimental Psychology, 49, 397-407.
- Baranski, J.V. & Petrusic, W.M. (1999). Realism of confidence in sensory discrimination. Perception and Psychophysics, 61, 1369-1383.
- Bjorkman, M. & Qvarsell, B. (1963). Confidence in S-E predictions. *Reports from the Psychological Laboratory. University of Stockholm*, No 155.
- Bjorkman, M., Juslin, P., & Winman, A. (1993). Realism of confidence in sensory discrimination: The underconfidence phenomenon. *Perception and Psychophysics*, 54, 75-81.
- Brandler, S. & Rammsayer, T.H. (1999). Differential practice effects in temporal discrimination: timing performance as a function of base duration and psychophysical procedure. In: *Fechner Day*, 99. Ed. P.R.Killeen & W.R.Uttal, Tempe: Arizona : Arizona State University, pp.228-233.
- Ferrel, W.R. (1995). A model for realism of confidence judgments: implications for underconfidence in sensory discrimination. *Perception and Psychophysics* 57, 246-254.
- Garriga-Trillo, A., Villarino, A., Gonzalez Labra, M. J. & Arnau, M. A. (1994). La calibraciyn de juicios psicofisicos: Estimaciyn de magnitudes (Calibration in psychophysical judgments: Magnitude estimation). *Psicothema*, 6 (3), 525-532. (In Spanish).
- Hellstrom, A. & Rammsayer, T. (2000). Time order effects in duration discrimination of noise bursts. In: *Fechner Day, 2000*. Ed. C.Bonnet. Strasbourg: France, Universite Louis Pasteur, pp.221-226.
- Juslin, P. & Ollson, H. (1997). Thurstonian and Brunswikian origins of uncertanty in judgments: A sampling model of confidence in sensory discrimination. *Psychological Review*, 104, 344-366.

- Olsson, H. & Winman, A. (1996). Underconfidence in sensory discrimination: The interaction between experimental setting and response strategies. *Perception and Psychophysics* 58, 374-382.
- Petrusic, W. M., & Baranski, J. V. (1998). Context, feedback, and the calibration and resolution of confidence in perceptual judgments. *American Journal of Psychology*, 110, 543-572.
- Rammsayer, T. H. & Wittkowski, K.M. (1990). Zeitfehler und Positionseffekt des Standardreizes bei der Diskrimination kurzer Zeitdauern. Archiv fur Psychologie, 142, 81-89.
- Ronis, D.L. & Yates, J.F. (1987). Components of probability judgement accuracy: individual consistency and effects of subject matter and assessment method. *Organizational Behavior and Human Decision Processes*, 40, 193-218.
- Skotnikova, I.G. (1994). (1). Accuracy, confidence-unconfidence and response times in unordered discrimination: subject-oriented research. (2). Confidence judgments specificity in same - different procedure (visual duration discrimination). In: *Fechner Day*, 1994. Ed. L.Ward. Vancouver, Canada: University of British Columbia, pp.208-219.
- Skotnikova, I. G. (2000). Calibration of confidence in different sensory tasks. In: *Fechner Day*, 2000. Ed. C.Bonnet. Strasbourg, France: Universite Louis Pasteur, pp.327-332.
- Stankov, L. (1998). Calibraton curves, scatterplotts and the distinction between general knowledge and perceptual tasks. *Learning and Individual Differences*, 10, 29-50.
- Whitcomb, K.M., Onkal, D. & Benson P.G. (1995). Probability judgment accuracy for general knowledge. Cross-national differences and assessment methods. *Journal of Behavioral Decision Mak ing*, 8, 51-67.
- Yates, J. F., Lee, J. & Bush, J. (1997). General knowledge overconfidence: Cross-national variations, response style, and "reality". Organizational Behaviour & Human Decision Processes, 70, 87-94.

ACKNOWLEDGEMENT: Preparation of this article was supported by the Deutsche Forschungsgemeinschaft.