Available online at www.sciencedirect.com

SCIENCE DIRECT®

International Journal of Psychophysiology 50 (2003) 101–115

INTERNATIONAL
JOURNAL OF
PSYCHOPHYSIOLOGY

www.elsevier.com/locate/ijpsycho

Combination of conflicting visual and non-visual information for estimating actively performed body turns in virtual reality

Simon Lambrey*, Alain Berthoz

LPPA, Collège de France-CNRS, 11 Place Marcelin Berthelot, Paris 75005, France

Received 14 October 2002; accepted 24 April 2003

Abstract

Whereas constant-weight linear models suffice for understanding many phenomena in the domain of perception and action, how the weights given to each sensory input are determined remains an open question. Notably, it has been suggested that weighting depends on the sensory context (e.g. the inconsistency between sensory signals) as well as on the subject. In the present study, the problem of non-linearity in multisensory interaction for estimating actively performed body turns was addressed at the level of group and individual data. Standing subjects viewed a virtual corridor in which forward movements were simulated at a constant linear velocity, and rotations were actually performed. Subjects were asked to learn the trajectory and then reproduce it from memory in total darkness. In the baseline condition, the relative amplitudes of visual and non-visual information for the performed rotations were the same, but were systematically manipulated in six 'sensory conflict' conditions. The subjects performed the task in these seven conditions 10 times (10 sessions), with a delay of at least 2 days between sessions. Five subjects placed more weight on visual than on non-visual information. The other 5 subjects placed more weight on non-visual than on visual information. Interestingly, the difference between 'visual' and 'non-visual' subjects in their use of conflicting information seemed to be accentuated by the fact of becoming aware of the sensory conflict. In all subjects, conflicting sensory inputs were combined in a linear way in order to estimate the angular displacements. However, signatures of non-linearity were detected when the data corresponding to the day on which subjects became aware of the conflict were considered in isolation. The present findings support the hypothesis that subjects used conflicting visual and non-visual information differently according to individual 'perceptive styles' (bottom-up processes) and that these 'perceptive styles' were made more observable by the subjects changing their perceptive strategy, i.e. re-weighting (top-down processes).

Keywords: Spatial memory; Sensory conflict; Human; Multisensory interaction; Virtual reality

---

*Corresponding author. Tel.: +33-1-44-27-13-86; fax: +33-1-44-27-13-82.
E-mail address: simon.lambrey@college-de-france.fr
(S. Lambrey).

0167-8760/03/$ - see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0167-8760(03)00127-2

1. Introduction

The sensory weighting model of multisensory integration consists of three processing layers (Zupan et al., 2002). Firstly, each sensor provides the central nervous system with information regarding a specific physical variable (S). Secondly, because the information available from different sensory systems is qualitatively different, sensory estimates are converted to intermediate estimates that share the same 'units', a process referred to as 'promotion' by Landy et al. (1995). This conversion is based on internal models of the relationships between sensory systems. Thirdly, because several sensory systems may provide information about the same physical variable, the final estimate ($\hat{S}$) is computed as a weighted average of all available intermediate sensory estimates as follows:

$$\hat{S}=\sum_i [w_i \times f_i(S)]$$

where the subscript i refers to a specific sensory modality, f is the operation by which the central nervous system estimates S and w is the weight given to a specific sensory estimate. The sensory weighting model has been extensively examined. How the weights are determined, however, remains an open question. According to the hypothesis of a maximum likelihood estimator (MLE), high weights are given to reliable cues and low weights to unreliable ones (see Jacobs, 2002 for a recent review): a cue is reliable if the distribution of inferences based on that cue has a relatively small variance, otherwise the cue is regarded as unreliable. Consistent with this hypothesis, a Kalman filter (or extended Kalman filter), which is an instance of MLE, has been proposed to model multisensory integration in, for instance, postural control (Gusev and Semenov, 1992; Kuo, 1995; van der Kooij et al., 1999; Kiemel et al., 2002), self-motion perception (Borah et al., 1988; Merfeld et al., 1993) or visual-haptic object perception (Ernst and Banks, 2002; Hillis et al., 2002). Some of these models view multisensory fusion as a constant-weight, linear process (e.g. Borah et al., 1988; Gusev and Semenov, 1992; Kuo, 1995; van der Kooij et al., 1999; Ernst and Banks, 2002; Kiemel et al., 2002). Linear weighting rules suffice for understanding many phenomena in the domain of perception and action, notably because all sensory systems normally provide congruent information about a specific physical variable (e.g. self-motion, posture or the shape of an object). Various studies, however, have reported evidence of non-linearity in multisensory interaction (e.g. Crowell et al., 1998; Mergner et al., 2000; Jeka et al., 2000; Oie et al., 2001; Lambrey et al., 2002; Oie et al., 2002; Hillis et al., 2002). For example, some authors, investigating the combination of touch and vision for postural control in humans, have suggested that multiple sensory inputs are dynamically re-weighted to maintain upright stance as sensory conditions change (Oie et al., 2002). In a study examining how human subjects combine visual and vestibular inputs for self-rotation perception, Mergner et al. (2000) reported findings that were incompatible with linear system predictions. The authors, therefore, described their results by a non-linear dynamic model in which the visual input can, in certain conditions, be suppressed by a visual-vestibular conflict mechanism. More recently, Lambrey et al. (2002) studied the effect of mismatched visual and non-visual information on the reproduction of actively performed body turns and discussed the hypothesis that one source of information may be dominant in estimating angular displacements, depending on the size of the sensory conflict and on the task to be performed (reproduction by body turns vs. drawing). These findings are consistent with non-linear cue conflict models that were proposed earlier (Zacharias and Young, 1981; Oman, 1982). They are also consistent with the robust estimator hypothesis (Hampel, 1974; Huber, 1981; Bulthoff and Mallot, 1988; Landy et al., 1995), which predict that as the discrepancy between conflicting cues increases beyond the range present in normal conditions, the weight given to the discrepant (outlier) cue will be increasingly low. In the event of major cue discrepancies, the outlier cue may even be completely disregarded, entailing the 'turning off' of the corresponding sensory modality. This phenomenon was referred to as 'vetoing behaviour' by Bulthoff and Mallot (1988).