Exp Brain Res (2009) 193:307–314
DOI 10.1007/s00221-008-1626-z

RESEARCH ARTICLE

Sensory dominance in combinations of audio, visual and haptic stimuli

David Hecht · Miriam Reiner

Received: 3 August 2008 / Accepted: 15 October 2008 / Published online: 5 November 2008

Abstract Participants presented with auditory, visual, or bi-sensory audio–visual stimuli in a speeded discrimination task, fail to respond to the auditory component of the bi-sensory trials significantly more often than they fail to respond to the visual component—a 'visual dominance' effect. The current study investigated further the sensory dominance phenomenon in all combinations of auditory, visual and haptic stimuli. We found a similar visual dominance effect also in bi-sensory trials of combined haptic–visual stimuli, but no bias towards either sensory modality in bi-sensory trials of haptic–auditory stimuli. When presented with tri-sensory trials of combined auditory–visual–haptic stimuli, participants made more errors of responding only to two corresponding sensory signals than errors of responding only to a single sensory modality, however, there were no biases towards either sensory modality (or sensory pairs) in the distribution of both types of errors (i.e. responding only to a single stimulus or to pairs of stimuli). These results suggest that while vision can dominate both the auditory and the haptic sensory modalities, it is limited to bi-sensory combinations in which the visual signal is combined with another single stimulus. However, in a tri-sensory combination when a visual signal is presented simultaneously with both the auditory and the haptic signals, the probability of missing two signals is much smaller than of missing only one signal and therefore the visual dominance disappears.

Keywords Sensory dominance · Visual dominance · Colavita effect · Modality appropriateness · Multi-sensory enhancement

Introduction

The way we perceive multi-sensory events reveals that our brain may not give equal weight to the information coming from the different sensory modalities. Rather, sometimes one sensory modality dominates the other. An everyday example of visual dominance over audition is the 'ventriloquism' effect experienced when watching television and movies, where the voices seem to emanate from the actors' lips rather than from the actual sound source (Pick et al. 1969; Howard and Templeton 1966; Alais and Burr 2004). Even more remarkable is the 'rubber-hand illusion' in which participants look at a *rubber hand* being stroked with a paintbrush while receiving a synchronous stroke on their own hidden *hand*. After a few minutes, when required to indicate the felt position of their hidden *hand* they point towards the *rubber hand* position, as if they experience the tactile stimuli arising from the rubber hand—an instance of visual dominance over proprioception and kinesthesia (Botvinick and Cohen 1998; Farnè et al. 2000; Pavani et al. 2000). Vision can also dominate smell and taste. A white wine surreptitiously colored with odorless red dye was described by enology students with language typically reserved for red wine and they avoided the use of white wine terms. Thus, when olfactory and visual information were incongruent, wine odor had minimal impact on olfactory discrimination and despite 'expertise' among participants, the visual contextual cue dominated (Morrot et al. 2001). In the same line, the perceived intensity of tastes and flavors can

D. Hecht (✉) · M. Reiner
The Touch Laboratory, Gutwirth Building,
Department of Education in Technology and Science,
Technion - Israel Institute of Technology, 32000 Haifa, Israel
e-mail: davidh@tx.technion.ac.il

change as a result of color-level manipulation (Roth et al. 1988; Delwiche 2004; Hoegg and Alba 2007).

In other circumstances, however, the other senses can dominate vision. A single flash of light accompanied by multiple auditory beeps is perceived as multiple flashes—an auditory dominance over vision (Shams et al. 2000, 2002). Similarly, participants presented with sequences of flashes, taps and beeps simultaneously were instructed to count the number of events presented in one modality (target) and to ignore the stimuli presented in the other modalities (background) as the number of events presented in the background sequence could differ from the target sequence. A comparison of participants' responses when the target was presented alone or with the background showed that vision was the most susceptible and the least efficient in biasing the other two senses. By contrast, audition was the least susceptible to background-evoked bias and the most efficient in biasing the other two senses (Bresciani et al. 2008). When participants touched the embossed tangible letters p, q, b, d, W, and M, while looking at them in an upright mirror that produced a vertical inversion of the letters and a visual inversion of the direction of finger movements, in a way that they touched the letter p but saw themselves in the mirror touching the letter b, most participants identified the letters relying on their touch and not on their vision (Heller 1992). In a gender discrimination study of ambiguous faces, the participants who inhaled androgen were more biased towards masculine judgments than the group exposed to estrogen (Kovacs et al. 2004).

A particular case of visual dominance was discovered by Colavita (1974). In speeded discrimination tasks, participants were asked to press a designated button when they detected a visual stimulus (flash), another button for an auditory stimulus (sound), and both buttons (or a third button) when both stimuli were presented together. In some of the bi-sensory trials participants failed and responded by pressing only one of the buttons as if only a single stimulus was presented. Remarkably, however, these erroneous responses were significantly biased towards the visual sensory modality, i.e., in the bi-sensory trials participants pressed only the visual button more often than they pressed only the auditory button (Colavita 1974; Colavita et al. 1976; Colavita and Weisberg 1979).

The 'Colavita effect' is a robust phenomenon that endured many experimental manipulations. For instance, the visual dominance persisted despite matching the subjective intensity of the two stimuli, or doubling the subjective intensity of the tone relative to that of the light (Colavita 1974). The effect was shown also regardless of whether uni-sensory auditory responses were slower than uni-sensory visual responses or vice versa (Koppen and Spence 2007a). Similarly, the effect was observed in simple detection tasks (e.g. responding to tone, flash, or both) as well as in more complex tasks—a go/no-go paradigm, in which predefined 'target' stimuli were interspersed in streams of distracter stimuli and participant responded only to the targets (Sinnett et al. 2007). The effect remained also when the probabilities of the uni- and bi-sensory trials, within experimental blocks, were varied (Koppen and Spence 2007a, c; Sinnett et al. 2007), although higher probabilities of bi-sensory stimuli reduced the magnitude of the effect. In the same line, the visual dominance persisted irrespective of the semantic congruence/incongruence between the auditory and the visual stimuli in the bi-sensory trials (Koppen et al. 2008).

The current study was designed to further explore the Colavita effect by investigating if there is a dominant modality also in bi-sensory combinations of visual and haptic, or auditory and haptic stimuli, and in tri-sensory combinations of auditory, visual and haptic stimuli.

Method

Participants

Twelve students participated in the experiment, six males and six females (mean age: 24.6 ± 2.6 years). Ten participants were right-handed, and two were left-handed according to the Edinburgh inventory (Oldfield 1971). All participants reported normal hearing and normal or corrected to normal vision and without any known tactile dysfunction. Participants gave their consent to be included in the study and were paid for their participation. They were unaware of the purpose of the experiment, except that it tested eye–hand coordination in different conditions. The experiment was carried out under the guidelines of the Technion's ethics committee.

Apparatus and stimuli

We used a virtual-reality (VR) touch-enabled computer interface capable of providing users with visual, auditory and haptic stimuli. The assembly included a computer screen that was tilted 45° and was reflected on a semitransparent horizontal mirror (Fig. 1). The participants viewed this reflection from above. A pen-like robotic arm (stylus) gripped and moved as in handwriting or drawing, was placed below the mirror surface. Full technical descriptions of this virtual haptic system are available at http://www.reachin.se and http://www.sensable.com.

The visual stimulus consisted of a thin, gray, horizontal line (length: 2.5 cm, width: 1 pixel). The auditory stimulus was a compound sound pattern of a horn (middle frequency: 11 kHz, 42 dB SPL) that was presented from two loudspeakers located at both sides of the stylus, approximately 35 cm

Fig. 1 Experimental setup. The visual display from the computer screen was reflected onto the horizontal mirror. Participants looked at the mirror while holding the pen-like stylus in their hand and positioning it at the center of their visual field. Two loudspeakers were placed in both sides of the stylus. In every trial the computer generated a uni-, bi- or tri-sensory stimulation, randomly, and participants were required to press the corresponding button(s)

from participants' ears. The haptic stimulus was a mechanical resisting force (0.35 Newton) delivered through the stylus, a pen-like robotic arm controlled by a programmable engine. The duration of all three stimuli was 600 ms. Since the Colavita effect is *maximal* when the auditory and visual stimuli are presented from the same spatial location (Koppen and Spence 2007b; Sinnett et al. 2007), in the current study all three sensory stimuli were presented at the same spatial location—the center of the workspace.

Procedure

Participants sat comfortably in front of the VR system, holding the stylus in their non-dominant hand and positioning its visual representation inside a circle (diameter: 1.5 cm) that was presented at the center of their visual field. They were instructed to stabilize their hand in that location by resting their stylus-holding arm on the table during the entire experimental session. Before a block of trials was initiated the graphic representation of the stylus disappeared off the screen. The dominant hand was placed on the response buttons device (SpaceMouse® Plus; http://www.3dconnexion.com), located on the dominant-hand side of the VR system. Participants were instructed to respond to each specific stimulus (auditory, visual, haptic), as soon as they detected it, by pressing a specific button designated for that stimulus. In the case of simultaneously occurring bi- or tri-sensory stimuli they were instructed to press the relevant two or three buttons. The response button designated for a given stimulus was constant for each participant during the entire experiment. However, the correspondence between a stimulus and its response button differed, in a balanced manner, between participants. For each trial, the computer registered the button(s) pressed as well as the response time (RT).

Trials were delivered in blocks with 3 min rest between blocks. Each block in the bi-sensory audio–visual, haptic–visual or audio–haptic conditions contained a randomly ordered mixture of 80 uni-sensory trials (40 of each uni-sensory stimulus) and 20 bi-sensory trials in which both stimuli occurred simultaneously. A tri-sensory block contained a randomly ordered mixture of 81 uni-sensory trials (27 of each uni-sensory stimulus) and 19 tri-sensory trials in which all three stimuli occurred simultaneously. The within-block ratios of ~80/20 (uni-/multi-sensory trials, respectively) were implemented in our study following the majority of previous studies on the Colavita effect that used this ratio, and as Koppen and Spence (2007c) found, the optimal Colavita effect occurs with a within-block majority of uni-sensory trials and a low ratio of bi-sensory trials. Each subject completed 10 blocks of each of the bi- and tri-sensory stimuli, totaling in 4,000 trials per participant. These trials were collected along four different days, i.e. each (bi- or tri-) sensory combination was tested in a different day. The order of the bi- and tri-sensory combinations was randomized and balanced across participants.

Prior to the experimental session, participants were trained briefly on their task before data recording began (about 20 trials in each stimuli combination). To ensure that the haptic signals were felt only *kinesthetically*, without additional *visual* cues of the hand movements, no direct visualization of the stylus holding hand was afforded by keeping the laboratory room darkened, and covering the participants' non-dominant hand with a black cloth.

Results

Distribution of errors

Overall, misses (0.02%) and inappropriate responses (i.e. pressing a visual button when an auditory signal was presented etc.; 0.73%) were distributed without significant difference among the visual, auditory and haptic modalities or their combinations. The errors of responding only to one

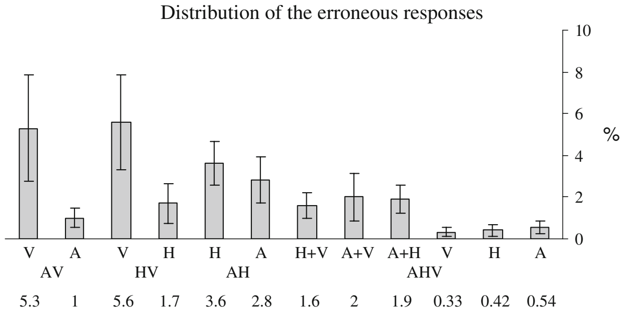


Fig. 2 Distribution of the erroneous responses in the bi- and tri-sensory combinations (mean ± standard deviation, pooled across participants). *Y* axis—error rate (in %), *X* axis; *upper line* the signals detected, *middle line* the sensory combination, *lower line* values of the *Y* axis

(or two) of the compound signals are summarized in Fig. 2. A paired *T* test showed that in the bi-sensory audio–visual trials participants made 5.3% errors of responding only with the visual button, significantly more than the 1% errors of responding only with the auditory button [$t_{(11)} = -3.06$, $P < 0.01$]. In the bi-sensory haptic–visual trials there were 5.6% errors of responding only with the visual button, significantly more than the 1.7% errors of responding only with the haptic button [$t_{(11)} = -3.24$, $P < 0.01$]. In the bi-sensory audio–haptic trials, participants erred in 2.8% of the trials by responding only with the auditory button, not significantly different from their 3.6% errors of responding only with the haptic button [Statistical power $(1 - \beta) > 96.6$].

In the tri-sensory audio–visual–haptic trials there were two types of errors. Responses to only a single sensory modality—auditory, visual or haptic—were 0.54, 0.33 and 0.42% respectively. Responses to only pairs of sensory modalities—auditory–visual, haptic–visual or auditory–haptic—were 2, 1.6 and 1.9% respectively. There were no significant differences between the auditory, visual and haptic modalities in the errors of responding only to a single sensory modality, or in the errors of responding only to pairs of sensory modalities [Statistical power $(1 - \beta) > 95.9$]. However, overall, participants made more errors of responding only to two sensory signals (5.5%) than errors of responding only to a single sensory modality—1.3% [$t_{(11)} = -5.27$, $P < 0.001$].

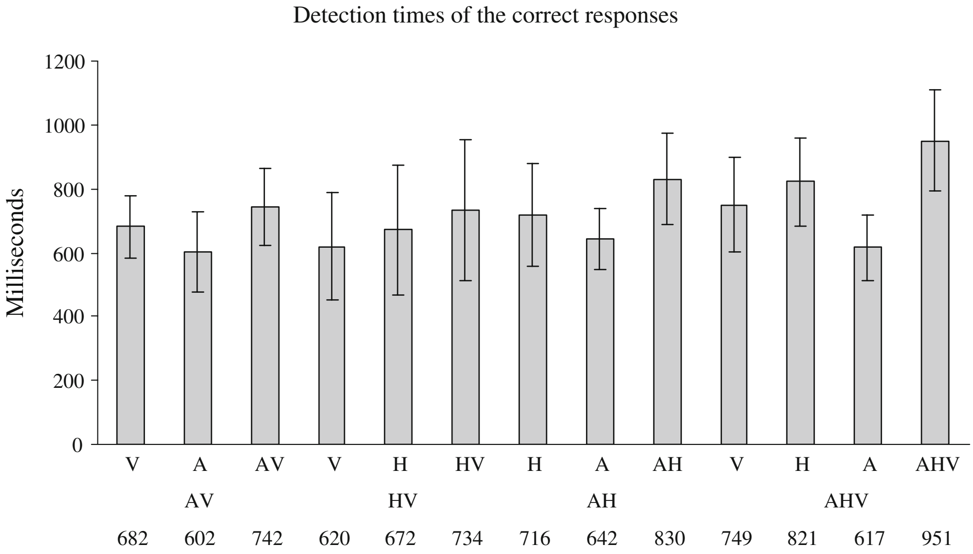
Within-participants analysis

The bias towards the visual modality in the bi-sensory combinations of audio–visual, and haptic–visual stimuli, was present in 10/12, 12/12 participants, respectively. In the bi-sensory combination of audio–haptic signals, five participants' errors were biased towards the auditory modality while the errors of the other five participants biased towards the haptic modality, the other two participants' errors were distributed equally between the auditory and haptic modalities. In the tri-sensory combinations, the trend of responding more to pairs of stimuli than to a single sensory modality occurred in all 12 participants.

Response times

Response times of the correct responses are summarized in Fig. 3. Four one-way repeated-measures ANOVA (1 for the tri-sensory combination, and 3 for the bi-sensory combination) with Bonferroni adjustment were conducted to analyze the RTs. The ANOVAs were followed by paired comparison analyses. In the blocks containing a mixture of uni-sensory audio, uni-sensory visual and bi-sensory

Fig. 3 Detection times of the correct responses in the uni- and multi-sensory combinations (mean ± SD, pooled across participants). *Y* axis—milliseconds, *X* axis; *upper line* the signals, *middle line* the sensory combination, *lower line* RT values of the *Y* axis



audio–visual trials there was an overall significant difference in RTs [$F_{(2,22)} = 14.45$, $P < 0.001$]. Paired comparison analyses showed that the difference between RTs to the uni-sensory visual signal (682 ± 98 ms) and RTs to the uni-sensory auditory signals (602 ± 127 ms) was significant [$t_{(11)} = -3.04$, $P < 0.01$). The RTs to the bi-sensory trials (742 ± 121 ms) were not significantly different from RTs to the uni-sensory visual signals, but significantly different from RTs to the uni-sensory auditory signals [$t_{(11)} = -7.63$, $P < 0.001$].

In the blocks containing a mixture of uni-sensory haptic, uni-sensory visual and bi-sensory haptic–visual trials there was an overall significant difference in RTs [$F_{(2,22)} = 20.28$, $P < 0.001$]. Paired comparison analyses showed that the difference between RTs to the uni-sensory visual and RTs to the uni-sensory haptic signals (620 ± 170 and 672 ± 203 ms, respectively) was significant [$t_{(11)} = 3.6$, $P < 0.005$). RTs to the bi-sensory trials (734 ± 219 ms) were significantly different from RTs to the uni-sensory visual signals [$t_{(11)} = -5.11$, $P < 0.001$], and different from RTs to the uni-sensory haptic signals [$t_{(11)} = -3.86$, $P < 0.005$].

An overall significant difference in RTs was found also in the blocks containing a mixture of uni-sensory auditory, uni-sensory haptic and bi-sensory audio–haptic trials [$F_{(2,22)} = 32.1$, $P < 0.001$]. Paired comparison analyses showed that the difference between RTs to the uni-sensory auditory and RTs to the uni-sensory haptic signals (642 ± 96 and 716 ± 161 ms, respectively) was significant [$t_{(11)} = -3.08$, $P < 0.01$]. RTs to the bi-sensory trials were 830 ± 143 ms, significantly different from RTs to the uni-sensory auditory signals [$t_{(11)} = -7.02$, $P < 0.001$], and different from RTs to the uni-sensory haptic signals [$t_{(11)} = -5.83$, $P < 0.001$].

In the blocks containing a mixture of uni-sensory audio, visual, or haptic trials and a tri-sensory combination of audio–visual–haptic trials, RTs to the uni-sensory auditory, visual and haptic trials were 617 ± 103, 749 ± 149, 821 ± 136 ms, respectively. An overall significant difference in RTs [$F_{(3,33)} = 35.12$, $P < 0.001$] was found. Paired comparison analyses showed that the difference between RTs to the uni-sensory auditory and the uni-sensory haptic signals was significant [$t_{(11)} = -6.35$, $P < 0.001$], as was the difference between RTs to the uni-sensory auditory and the uni-sensory visual signals [$t_{(11)} = 4.6$, $P < 0.001$], and the difference between RTs to the uni-sensory visual and the uni-sensory haptic signals [$t_{(11)} = -2.34$, $P < 0.05$]. RTs to the tri-sensory audio–visual–haptic trials were 951 ± 160 ms, significantly different from RTs to the uni-sensory auditory signals [$t_{(11)} = -8.89$, $P < 0.001$], from RTs to the uni-sensory visual signals [$t_{(11)} = -7.39$, $P < 0.001$), and from RTs to the uni-sensory haptic signals [$t_{(11)} = -3.16$, $P < 0.01$].

Discussion

The results of the current study replicated Colavita's findings (Colavita 1974; Colavita et al. 1976; Colavita and Weisberg 1979) that in a compound of auditory and visual signals, it is more likely for the auditory signal to be unnoticed than for the visual signal. Furthermore, the current study extends the 'visual dominance' phenomenon by showing that in a compound of haptic and visual signals, it is also more likely to be unaware of the presence of the haptic signal than of not noticing the visual signal. The fact that there was no significant bias towards either sensory modality in compounds of haptic and auditory signals, suggests that while there is a prepotency and dominance of the visual system over the auditory and the somatosensory systems, there is no natural hierarchy between the auditory and the somatosensory systems. These conclusions are further supported by the convergence of the group-averages and the within-participants analyses, indicating that the dominance of the visual system was characteristic of most of the individual participants' performances in the audio–visual and haptic–visual blocks. However, in the audio–haptic blocks there were equal numbers of participants, whose errors were biased towards the auditory or towards the haptic sensory modalities, suggesting no dominance of one sensory modality over the other for the auditory and haptic systems.

The results of the current study also show that the occurrence of Colavita's visual dominance effect is limited to bi-sensory combinations when a visual signal is synchronized with an auditory or a haptic signal, whereas in tri-sensory combinations of audio–visual–haptic signals there was no bias towards vision in both types of errors (i.e. in the errors of responding only to one sensory signal there was no significant difference between the senses, and also in the errors of responding only to two sensory signals there was no bias towards responses that contained a visual element). The significant tendency, in the tri-sensory blocks, towards more errors of responding only to two sensory signals than errors of responding only to a single sensory signal, is very reasonable since by responding only to two signals, a *single* error is being made of not noticing a single cue, whereas by responding only to one signal, *two* errors are being made—not noticing two cues. The probability of missing two signals is less than of missing only one signal.

In the Colavita paradigm (and its current study extensions) participants were engaged in multi-tasking that required allocating attention and working memory resources in multiple channels simultaneously. The multiple resources model (Wickens 2002, 2008) describe multitasking on a 4-dimensional scheme in which tasks may differ on the: (1) stage—perception/cognition versus response, (2) processing code—spatial versus verbal, (3) sensory modality—visual versus auditory etc., (4) for a visual task—focal versus peripheral. This model predicts greater interference when different time-sharing tasks are carried out within the same dimension (e.g. *looking* for directions while driving is more demanding than *listening* to directions etc.). In the current study, the overall error rate in the uni-sensory trials was 0.75% (combined misses and inappropriate responses), much lower than the approximately ~6% [combined errors of responding only to part(s) of the compound signals] in the multisensory trials where participants were required to make multiple responses. The multiple resources model may explain the relatively larger proportion of errors in the multisensory trials, despite the utilization of different sensory modalities, as a result of multiple tasks time-sharing the same *stages* (initially, detecting the signals, and later executing the motor response) and the same *response mode* [manual—pressing button(s)]. Nevertheless, the main effect—that errors are not distributed equally and that it is more likely to be unaware of an auditory or a haptic signal than to be unaware of a visual signal—is still unexplained unless some degree of visual dominance over the auditory and haptic systems is assumed.

Sensory dominance

When presented with incongruous cues from different sensory modalities, the ‘modality appropriateness hypothesis’ (Welch and Warren 1980) postulates that the sensory system that has the greatest precision for a given task will dominate perception. Thus, the visual system dominates in *spatial* tasks where it has a greater acuity, while *temporal* tasks are dominated by the auditory system with its superior temporal resolution (Welch and Warren 1980; Recanzone 2003). Despite the ability of the modality appropriateness hypothesis to explain several phenomena such as the ventriloquism effect, the rubber-hand illusion (both are spatial tasks where vision dominates) and auditory influences on vision in temporal tasks (e.g. Shams et al. 2000, 2002; Bresciani et al. 2008), this hypothesis is limited in scope and insufficient (see also Shams et al. 2004). First, besides spatial and temporal tasks, it does not provide clear predictions for other tasks. Second, it cannot account for the aforementioned studies in which a purely gustatory task—rating the intensity of tastes and flavors—could be dominated by vision (through manipulations of food colorants’ level; Roth et al. 1988; Delwiche 2004; Hoegg and Alba 2007), and an olfactory task—describing the hedonic qualities of a wine by means of smelling it—was dominated primarily by vision (Morrot et al. 2001). Third, it is irrelevant in explaining the Colavita effect, since in Colavita’s paradigm the task was simply to detect the occurrence of the stimuli without any localizations or temporal judgments requests, and the findings clearly indicated that even in the initial detection of the stimuli it was more likely for the auditory and haptic signals to be unnoticed than for the visual signal.

A more sophisticated approach, elaborating the modality appropriateness’s concept with Bayesian statistics principles was recently proposed. It is based on the notion that each sensory modality by itself provides the CNS with imperfect and variable sensory inputs. According to Bayesian inference principles the imperfect estimate obtained from one sensory input can be improved by taking into account the probabilities of signals from another sensory modality. Thus, our brain often minimizes the uncertainty of imperfect and noisy sensory inputs by combining probabilities of multiple sensory signals to refine sensory estimates. In these optimal estimates, prior experiences are also taken into account and the nervous system gives more weight to the less variable estimate, thus in ambiguous or incongruous conditions, the sensory modality that affords the most precise estimate at that moment contributes to perception more than the other sensory modalities do (Ernst and Banks 2002; Ernst and Bülthoff 2004; Alais and Burr 2004; Gepshtein et al. 2005; Körding 2007; Körding et al. 2007). This Bayesian inference approach that does not restrict itself to a rigid linkage of particular sensory systems with specific tasks (as the original ‘modality appropriateness’s proposition) has a better explanatory power for sensory dominance phenomena. For instance, although rating the intensity of tastes and flavors is primarily a gustatory task, if the visual cue (color) is more salient and prior experience associates taste intensity with color level (e.g. a ripe fruit vs. an almost-ripe fruit etc.) the CNS may prefer the visual cue over the gustatory cue which may had a poorer resolution in that particular situation (e.g. Roth et al. 1988; Delwiche 2004; Hoegg and Alba 2007). Likewise, the description of wine qualities may be dominated by its color, not its aroma, if the smell does not correspond with previous knowledge about wines’ color–aroma relationships (e.g. Morrot et al. 2001). This may be especially applicable if participants do not swallow the wine and only smell it, so the olfactory cues are isolated from the regularly-accompanied gustatory cues (and therefore less prominent). In the same line, when the visual cue is vague, the brain may utilize the co-occurring sex hormone-like compounds as a better cue for determining the gender of a morphed face (e.g. Kovacs et al. 2004).

This Bayesian optimal inference approach can explain also some aspects of the Colavita effect and its current study’s extensions. For instance, it may explain how and why participants failed in ~6% of the trials to detect both components of the bi-sensory compounds, in terms of inherent noise which may have caused some variance in the neural transmission and consequently in the detectability of the signals. Similarly, the tendency in the tri-sensory blocks towards more errors of responding only to two sensory signals than errors of responding only to a single sensory signal, can be explained in terms of lower probability for missing two signals compared to the probability of missing only one signal. In the same line, it may explain why the visual dominance is most likely to be found in bi-sensory combinations, but not in a tri-sensory combination when a visual signal is presented simultaneously with both the auditory and the haptic signals, since the probability of missing two signals is less than of missing one signal. Nevertheless, the core findings of Colavita and the current study—that it is more likely to be unaware of an auditory or a haptic signal than to be unaware of a visual signal—is still unexplained since noise and variance in signal-transduction should have been distributed equally without significant differences among the corresponding sensory modalities, especially in large samples (1,000 trials per participant for each combination in the current study). That is, unless some degree of visual prepotency and dominance over the auditory and haptic systems is assumed, at least for the initial detection of the signals.

The dominance of the visual system that was observed in the current study is not unique to humans and it had been observed also in pigeons (Randich et al. 1978) and in rats (Miller 1973; Meltzer and Masaki 1973; Bushnell and Weiss 1977). However, an opposite pattern was found in cats where they responded more to the auditory signals than to the visual signals (Jane et al. 1965). Thus, the hierarchy between the sensory modalities may be specific for every species, depending on its environmental and neural properties, as well as on evolutionary adaptive strategies. Moreover, even in humans, visual dominance is not congenital. On the contrary, the development of the fetus's brain is asynchronous and the auditory system precedes the visual system structurally and functionally (Bronson 1982; Lewkowicz 1988a; Liu et al. 2007). Studies with infants exposed to an audio–visual compound stimulus showed that 6-month-old infants discriminated changes in the temporal characteristics of the auditory component but never discriminated such changes in the visual component. However, 10-month-old infants could, under certain conditions, discriminate also temporal changes in the visual component (Lewkowicz 1988a, b). Based on these findings, Lewkowicz (1988b) proposed that the beginning of the developmental shift, in human infants, from 'auditory dominance' towards 'visual dominance' occurs somewhere during the time span between 6 and 10 months of age.

Regarding the response times, an interesting pattern was revealed here. Typically, compounds of multi-sensory signals are detected faster than when the same signals are presented separately—a phenomenon known as *multisensory enhancement* (Hershenson 1962; Doyle and Snowden 2001; Forster et al. 2002; Fort et al. 2002; Hecht et al. 2008a, b). In this study, however, when participants responded correctly, their RT for the bi- and tri-sensory trials were *slower* than their RT to the corresponding uni-sensory trials. The explanation of this apparent discrepancy lies in the differences in task requirements. Whereas in the multisensory enhancement studies a *redundant signal paradigm* was used in which participants were asked to respond in the multi-sensory trials 'as soon as they detected *any* of the signals', by pressing the *same* button in the uni- and multi-sensory trials, in the current study, after detecting the presence of the signal(s) participants needed to discriminate the signals according to their sensory modalities and to choose from *different* response buttons the appropriate button or buttons. Consequently, responses to uni-sensory trials were shorter as participants needed to make only one decision, whereas in the bi- and tri-sensory trials they were required to make two (or 3) separate decisions—one for each sensory modality—and the additional cognitive process resulted in longer RTs. A similar effect—an increase in the number of decisions-to-be-made entails increase in RTs—was reported by Hyman (1953) where RTs were prolonged in accord with the number of response alternatives.

In conclusion, the results of the current study show that: (1) Vision can dominate not only the auditory but also the haptic sensory modality. (2) There is no dominance of the auditory sensory modality over the haptic sensory modality or vice versa. (3) In a tri-sensory combination of audio–visual–haptic signals participants tend to err and respond only to two signals (out of 3) more than they err by responding only to a single stimulus. (4) Since the probability of missing two signals is less than of missing one signal, the visual dominance is most likely to be found in bi-sensory combinations, but not in a tri-sensory combination when a visual signal is presented simultaneously with both the auditory and the haptic signals. Future studies may further broaden our knowledge on the dynamics of vision, audition and the haptic sensory modalities when combined, in the context of simple detection task, with the chemical senses—olfaction and gustation.

Acknowledgments This research was funded by the EU research project PRESENCCIA—Presence: Research Encompassing Sensory Enhancement, Neuroscience, Cerebral–Computer Interfaces and Applications. We thank Mr. Gad Halevy for programming the computer for the experiment.

References

Alais D, Burr D (2004) The ventriloquist effect results from near-optimal bimodal integration. Curr Biol 14(3):257–262

Botvinick M, Cohen J (1998) Rubber hand 'feel' touch that eyes see. Nature 391(6669):756

Bresciani JP, Dammeier F, Ernst MO (2008) Tri-modal integration of visual, tactile and auditory signals for the perception of sequences of events. Brain Res Bull 75(6):753–760

Bronson GW (1982) Structure, status, and characteristics of the nervous system at birth. In: Stratton P (ed) Psychobiology of the human newborn. Wiley, New York, pp 99–118

Bushnell MC, Weiss SJ (1977) The effect of reinforcement differences on choice and response distribution during stimulus compounding. J Exp Anal Behav 27(2):351–362

Colavita FB (1974) Human sensory dominance. Percept Psychophys 16(2):409–412

Colavita FB, Weisberg D (1979) A further investigation of visual dominance. Percept Psychophys 25(4):345–347

Colavita FB, Tomko R, Weisberg D (1976) Visual prepotency and eye orientation. Bull Psychon Soc 8:25–26

Delwiche J (2004) The impact of perceptual interactions on perceived flavor. Food Qual Prefer 15(2):137–146

Doyle MC, Snowden RJ (2001) Identification of visual stimuli is improved by accompanying auditory stimuli: the role of eye movements and sound location. Perception 30(7):795–810

Ernst MO, Bülthoff HH (2004) Merging the senses into a robust percept. Trends Cogn Sci 8(4):162–169

Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. Nature 415(6870):429–433

Farnè A, Pavani F, Meneghello F, Làdavas E (2000) Left tactile extinction following visual stimulation of a rubber hand. Brain 123(11):2350–2360

Forster B, Cavina-Pratesi C, Aglioti S, Berlucchi G (2002) Redundant target effect and intersensory facilitation from visual–tactile interactions in simple reaction time. Exp Brain Res 143(4):480–487

Fort A, Delpuech C, Pernier J, Giard MH (2002) Dynamics of cortico-subcortical cross-modal operations involved in audio–visual object detection in humans. Cereb Cortex 12(10):1031–1039

Gepshtein S, Burge J, Ernst MO, Banks MS (2005) The combination of vision and touch depends on spatial proximity. J Vis 5(11):1013–1023

Hecht D, Reiner M, Karni A (2008a) Enhancement of response times to bi- and tri-modal sensory stimuli during active movements. Exp Brain Res 185(4):655–665

Hecht D, Reiner M, Karni A (2008b) Multisensory enhancement: gains in choice and in simple response times. Exp Brain Res 189(2):133–143

Heller MA (1992) Haptic dominance in form perception: vision versus proprioception. Perception 21(5):655–660

Hershenson M (1962) Reaction time as a measure of intersensory facilitation. J Exp Psychol 63:289–293

Hoegg J, Alba JW (2007) Taste perception: more than meets the tongue. J Consum Res 33(4):490–498

Howard IP, Templeton WB (1966) Human spatial orientation. Wiley, New York

Hyman R (1953) Stimulus information as a determinant of reaction time. J Exp Psychol 45(3):188–196

Jane JA, Masterton RB, Diamond IT (1965) The function of the tectum for attention to auditory stimuli in the cat. J Comp Neurol 125(2):165–191

Koppen C, Spence C (2007a) Seeing the light: exploring the Colavita visual dominance effect. Exp Brain Res 180(4):737–754

Koppen C, Spence C (2007b) Spatial coincidence modulates the Colavita visual dominance effect. Neurosci Lett 417(2):107–111

Koppen C, Spence C (2007c) Assessing the role of stimulus probability on the Colavita visual dominance effect. Neurosci Lett 418(3):266–271

Koppen C, Alsius A, Spence C (2008) Semantic congruency and the Colavita visual dominance effect. Exp Brain Res 184(4):533–546

Körding KP (2007) Decision theory: what “should” the nervous system do? Science 318(5850):606–610

Körding KP, Beierholm U, Ma WJ, Quartz S, Tenenbaum JB, Shams L (2007) Causal inference in multisensory perception. PLoS ONE 2(9):e943

Kovacs G, Gulyas B, Savic I, Perrett DI, Cornwell RE, Little AC, Jones BC, Burt DM, Gal V, Vidnyanszky Z (2004) Smelling human sex hormone-like compounds affects face gender judgment of men. Neuroreport 15(8):1275–1277

Lewkowicz DJ (1988a) Sensory dominance in infants: I. Six-month-old infants’ response to auditory–visual compounds. Dev Psychol 24(2):155–171

Lewkowicz DJ (1988b) Sensory dominance in infants: II. Ten-month-old infants’ response to auditory–visual compounds. Dev Psychol 24(2):171–182

Liu WF, Laudert S, Perkins B, MacMillan-York E, Martin S, Graven S (2007) The development of potentially better practices to support the neurodevelopment of infants in the NICU. J Perinatol 27(Suppl 2):S48–S74

Meltzer D, Masaki MA (1973) Measures of stimulus control and stimulus dominance. Bull Psychon Soc 1:28–30

Miller L (1973) Compounding of discriminative stimuli that maintain responding on separate response levers. J Exp Anal Behav 20(1):57–69

Morrot G, Brochet F, Dubourdieu D (2001) The color of odors. Brain Lang 79(2):309–320

Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9(1):97–113

Pavani F, Spence C, Driver J (2000) Visual capture of touch: out-of-the-body experiences with rubber gloves. Psychol Sci 11(5):353–359

Pick HL, Warren DH, Hay JC (1969) Sensory conflict in judgments of spatial direction. Percept Psychophys 6:203–205

Randich A, Klein RM, LoLordo VM (1978) Visual dominance in the pigeon. J Exp Anal Behav 30(2):129–137

Recanzone GH (2003) Auditory influences on visual temporal rate perception. J Neurophysiol 89(2):1078–1093

Roth HA, Radle LJ, Gifford SR, Clydesdale FM (1988) Psychophysical relationships between perceived sweetness and color in lemon- and lime-flavored drinks. J Food Sci 53(4):1116–1119

Shams L, Kamitani Y, Shimojo S (2000) Illusions: what you see is what you hear. Nature 408(6814):788

Shams L, Kamitani Y, Shimojo S (2002) Visual illusion induced by sound. Brain Res Cogn Brain Res 14(1):147–152

Shams L, Kamitani Y, Shimojo S (2004) Modulation of visual perception by sound. In: Calvert G, Spence C, Stein B (eds) The handbook of multisensory processes. MIT Press, Cambridge, pp 27–33

Sinnett S, Spence C, Soto-Faraco S (2007) Visual dominance and attention: the Colavita effect revisited. Percept Psychophys 69(5):673–686

Welch RB, Warren DH (1980) Immediate perceptual response to intersensory discrepancy. Psychol Bull 88:638–667

Wickens CD (2002) Multiple resources and performance prediction. Theor Issues Ergon Sci 3(2):159–177

Wickens CD (2008) Multiple resources and mental workload. Hum Factors 50(3):449–455