Crossmodal correspondences between visual, olfactory and auditory information
Our senses take in a large amount of information, information that sometimes is congruent across sensory modalities. Crossmodal correspondences are the study of how this information across modalities is integrated by the brain, across which dimensions the correspondences exists, and how it affect us. In the present paper four experiments were conducted, in which potential crossmodal correspondences between audition, vision and olfaction were investigated. It was hypothesized that crossmodal correspondences between olfaction, vision and audition exist along different dimensions. The results showed significant correlations between olfaction and audition when volume varies, i.e., a high volume is associated to a high concentration of an odor, and a low volume is associated to a low concentration of an odor, and vice versa. Furthermore, existing correspondences between vision and audition is reconfirmed. In conclusion, the results provide support to the notion that crossmodal correspondences exists between all sensory modalities, although along different dimensions.

Traditionally the human sensory systems (e.g., vision, audition, and olfaction) have been viewed as separated entities. However, in many everyday situations, the information we take in about a single object often comes from many different senses, and becomes an integrated part of the perceptual experience. Sometimes the information from the senses can lead to distractions, as when trying to read in a café and your attention is drawn towards your neighbor’s loud conversation. Other times, for example while in a conversation, the input from vision (i.e., reading the lips and facial expressions) can modulate the speech perception, and at times even fully alter it, as in the McGurk-effect (McGurk & MacDonald, 1976). Eating is another example of the same mechanism; how the food looks (Spence et al., 2010), the texture and smell of it (Bult, de Wijk & Hummel, 2007), the temperature of it (Cruz & Green, 2000), the interactions of sweet, sour, umami, etc., and even the sound while chewing (Zampini & Spence, 2010) affects our taste-experience. The object of this paper is to explore if dimensions of concentrations in olfactory, visual and auditory information show any correspondence between each other.

Above are examples of what is known as crossmodal correspondences, which refers to a “compatibility effect between attributes or dimensions of a stimulus (i.e., an object or event) in different sensory modalities (be they redundant or not)” (Spence, 2011). One of the most widely held views on when, and why, crossmodal correspondences occur, is the assumption of unity (e.g., Vroomen & Keetels, 2010). When different modalities process information that shares some amodal property, for example temporal or spatial properties, the more likely it is that the brain treats them as coming from the same
source. The existence of crossmodal correspondences has been known for many years (Köhler, 1929; Sapir, 1929; see Spence, 2011, for a review), although at first, interest was largely focused on the subject of sound symbolism (see Nuckolls, 2003, for a review), that is, the idea that vocal sounds have intrinsic meaning. In Sapir’s (1929) experiment participants were instructed to associate two objects (i.e., a big or a small circle) to the nonsense words “mal” or “mil”. It was hypothesized that the speech sound of /a/ would correspond to the bigger circle, while /i/ would correspond to the smaller circle. Results confirmed the hypothesis that individuals tend to associate “mal” to a large object and “mil” to a small object. In a similar experiment Köhler (1929) found that people tend to associate a rounded shape with the nonsense-word “Baluma”, and a sharp, pointy-edged shape to the nonsense-word “Takete”. This initial research led to what is now known as the “Kiki/Bouba”-effect (Ramachandran & Hubbard, 2001). Ramachandran and Hubbard (2001) demonstrated that between 95% and 98% of the population agreed on which object is Kiki and which is Bouba, when presented with an edgy and a round shape. Also, the effect is showed to be cross-cultural. Other basic findings include, for instance, matching of visual brightness and loudness of sounds (Stevens & Marks, 1965) and matching between high-pitched sounds and a high elevation in space (Mudd, 1963), to name a few. During the past decades the topic of crossmodal correspondences has gained in interest. Crossmodal correspondences have been found between many different sensory modalities, for example: between audition and touch (Yau et al., 2009), olfaction and touch (Dematté et al., 2006a), vision and touch (Martino & Marks, 2000), auditory pitch and smell (Belkin et al., 1997), color and olfaction (Dematté et al., 2006b; Kemp & Gilbert, 1996), shapes and olfaction (Seo et al., 2010), color and taste (Spence et al., 2010), nonsense-words/shapes and taste (Spence & Gallace, 2011), and taste and musical notes (Crisinel & Spence, 2010).

Some authors argue that the mechanisms underlying crossmodal correspondences are in fact a weaker form of synaesthesia (Martino & Marks, 2001; Sagiv & Ward, 2006; Ward, Huckstep, & Tsakanikos, 2006), which is a perceptual phenomena in which stimulation of one perceptual modality evokes an automatic perception in a second sensory modality (Baron-Cohen & Harrison, 1997). For example, a musical note may trigger the perception of colors or shapes. Martino and Marks (2001) argued that all individuals are synaesthetes, but that the degree of synaesthesia ranges from weak to strong across individuals. Furthermore, Ramachandran and Hubbard (2001) suggested that synaesthesia may have triggered the evolution of language. Although the issue is yet unresolved, Spence (2011), among others (see also Elias et al., 2003), argued against this view and suggested that there is resemblance between synaesthesia and crossmodal correspondences, though there are more differences than similarities between the two. Whichever is the correct interpretation of the current research, both research areas complement each other in the understanding of human perception.

While some of the studies on crossmodal correspondence have been focusing on finding relationships between features of different stimuli, others have been looking into how crossmodal correspondences affect our information processing, often by using the speeded classification task (e.g., Gallace & Spence, 2006; Keetels & Vroomen, 2010). In this experimental paradigm, participants are presented with, for example, a visual object and a sound, and are told to decide the size of the visual object while ignoring the sound. When a visual object and a sound are congruent along a specific dimension, e.g.,
size (e.g., big circle/low-pitched sound), individual’s response latencies for the visual stimuli decreases. Thus, crossmodal correspondences depend on an interaction between two or more sensory systems, which also affect our behavior.

Spence (2011) suggested that three types of crossmodal correspondences exist: Structural, which refers to the neural processing of sensory information and depends on the cognitive systems. Statistical, which imply that correspondences are found in nature and are learned accordingly. For example, large objects make loud noises; the frequency of a certain object is related to its mass (Grassi, 2005), and so forth. These natural occurring associations are more likely to occur universally, i.e., in most people and cross culturally. Semantically mediated, that is when two dimension match with regard to the semantic meaning of the stimuli.

Most of the current research has been concerned with the interactions between vision and hearing (see Spence, 2011, for a review) while research concerning odor and its relations to other modalities is not as extensive. For instance, Belkin et al., (1997) investigated whether qualitatively different odors corresponded with different pitch of a sinus tone, finding that this was the case. Furthermore, Kemp and Gilbert (1997) found correspondences between visual lightness and odor intensity, using three different concentrations of different odors. Moreover, in an experiment involving white and red wine, it was shown that by coloring white wine red with an odorless colorant, the olfactory information was discounted for (Morrot et al., 2001).

In the present paper four experiments are presented. The overall aim was to study potential correspondences between olfaction, vision and audition, thus trying to replicate and extend previous work on correspondences between audition and vision and odor and vision. We also wanted to investigate the relationship between concentrations of odor and audition, manipulating the volume of the auditory stimulus. It was hypothesized that (i) a low concentration of an odor would be associated with a high frequency/low volume of a sound and a lighter shade of grey, (ii) a high concentration of an odor would be associated with a low frequency/high volume of a sound and a darker shade of grey, (iii) a low frequency/high volume of a sound would be associated with a darker shade of grey and a high concentration of an odor, (iv) a high frequency/low volume of a sound would be associated with a lighter shade of grey and a low concentration of an odor, (v) a lighter shade of grey would be associated with a high frequency/low volume of a sound and a low concentration of an odor, and (vi) a darker shade of grey would be associated with a low frequency/high volume of a sound and a high concentration of an odor. The present study is novel in that all correspondences are tested in the same experiment, the dimensions of concentration are varied for all stimuli, and both the frequency and the volume of sounds are included.

General Method

Participants
Participants were recruited through personal contact and by ads posted on the university campus. All participants were informed that the test would involve smelling different substances, seeing different pictures and hearing different sounds. They were also
informed that their participation was voluntary, and that they could discontinue the test at any time.

Procedure
The experiment took place in a standardized laboratory. Each subject was presented with a brief explanation of the experiment, followed by a test-trial where they were familiarized with the different stimuli. In Experiment 1, the presentation was controlled by the experiment leader, and the constant auditory stimuli only lasted for 14 seconds. This was changed for the remaining experiments (Experiment 2-4), where the presentation was self-paced, i.e., the participants themselves clicked to see the next presentation. When clicking on a keyboard, the presentation started, and one constant stimulus was presented. For example, if the picture containing the darkest color was constant, it was shown on the screen while the five different sounds, or the five different odors, were presented at the same time (see Figure 1). The same goes for the sounds, a constant sound was presented while the five odors or the five pictures were presented on the screen. After each presentation there was a short break where the subject filled out a form choosing which one of the five best corresponded to the constant. When an odor was the constant, the subject was given the jar/Sniffin' stick, and was told to take one sniff (approximately 2 seconds of sniffing) per matching sound or picture. In Experiment 1, this part was automatic, which was later changed for Experiment 2 to 4; it was then self-paced (i.e., the subject held the jar containing the odor in one hand, and changed matching stimuli on click). Each matching stimuli was presented for 2 seconds, while the constant stimuli at the same time was being presented to the subject. When odors were the matching stimuli, the assistant presented them to the subject, keeping the number of the jars hidden from the subject’s view. Each constant stimulus was presented twice to match against the two other categories. All presentations were randomized.

Figure 1. Example of the experimental procedure. Each big square represents a constant slide on the computer screen.
**Experiment 1**
The aim of Experiment 1 was to test the hypothesis that there exists corresponding dimensions between olfactory, vision and audition regarding concentration.

**Method**

**Participants**
Fifteen students (10 females and 5 males; age range 21-35 years; mean age $M=27.13$, $SD=4$) at the Department of Psychology, Stockholm University, participated in Experiment 1 for course credit.

**Materials**
The olfactory materials used in Experiment 1 consisted of the following five concentrations of whiskey: 4%, 8%, 16%, 32%, and 64% whiskey. All odors were kept in nontranslucent glass jars to prevent visual inspection. Each jar contained 10 ml of whiskey. Tap water was used as solvent. Five Sinus-tones in 533Hz, 1134Hz, 1400Hz, 1534Hz and 1600Hz was created using the program NCH Tone Generator. As visual stimuli, a black picture in which the transparency was changed to 96%, 92%, 84%, 68% and 36% was used. The presentation was carried out using Microsoft Power Point. It was constructed in the following way: 30 Power Point-slides were used, where each stimulus was presented twice (3 categories x 5 concentrations x 2) as a constant. For each constant, 5 slides were presented from one of the remaining two categories (i.e., if the darkest picture was the constant, then all odors, or all sounds, would be presented). Four different randomized Power Point-presentations were used.

**Results and Discussion**
Figure 2 show the significant and non-significant correlations between stimulus-categories. Data was analyzed using Spearman’s rank correlation coefficient for non-parametric tests. First, correlations for each individual were calculated. Next, a mean correlation across all subjects was calculated. The mean correlation was then subjected to an independent one sample t-test. This yielded a significant negative correlation between vision and audition ($r = -0.427$, $t_{14} = -3.129$, $p < .01$), such that a darker grey picture corresponded to a lower pitch and a lighter grey picture corresponded to a higher pitch; and a negative significant correlation between audition and vision ($r = -0.330$, $t_{14} = -2.244$, $p < .05$), a low pitch corresponded to a darker grey picture and a high pitch corresponded to a lighter grey picture. All other correlations were non-significant. These results provide partial support for the present hypotheses and replicate previous research. Thus, even though no significant correlations where found for olfaction in either direction, the significant correlations between vision and audition, in both directions, may be interpreted as further evidence for the hypothesis that crossmodal correspondence arise when information from two sensory systems are matched along a common dimension.
Figure 2. Correlations between audition, olfaction and vision, Experiment 1.

Experiment 2
The aim of this experiment was to follow up on Experiment 1, with some changes in the design of the experiment. The experiment was self-paced, and the olfactory stimuli were changed to n-butanol.

Method

Participants
Fifteen students (11 females and 4 males; age range 21-33 years, mean age $M=25.2$, $SD=3.63$) at the Department of Psychology, Stockholm University, participated in Experiment 2 for course credit.

Materials
In Experiment 2, the olfactory stimuli consisted of the following concentrations of n-butanol (i.e., custom specified and factory calibrated Sniffin' Sticks): 4%, 8%, 16%, 32%, and 64%. These concentrations were the same as for the whiskey concentrations in Experiment 1. Apart from these changes everything else remained the same as in Experiment 1.

Results and Discussion
Figure 3 show the significant and non-significant correlations between stimulus-categories. The same analyses as in Experiment 1 were conducted. The results showed a negative significant correlation between audition and vision ($r=-0.374$, $t_{14}=-3.977$, $p<.01$), such that a low pitch corresponded to a darker gray picture and a high pitch corresponded to a lighter gray picture. All other correlations were non-significant. The
results in Experiment 2 are similar to the ones in Experiment 1, although one of the significant correlations, that between vision and audition, disappeared. This could be because of the change of design to self-paced, but that seems unlikely. Neither the visual nor the auditory stimuli was changed, only the olfactory stimuli were changed, from whiskey in Experiment 1 to Sniffin' Sticks in Experiment 2. The most likely explanation would be that this happened by chance. The correlation between olfactory and visual/audition remain non-significant, which is in line with Experiment 1.

Figure 3. Correlations between audition, olfaction and vision, Experiment 2.

**Experiment 3**
In Experiment 3, some further changes were made to the experiment. The sound was changed from frequency to volume to test the hypothesis that dimensions of sounds are correlated to dimensions of concentration of an odor.

**Method**

**Participants**
Fourteen students (9 females and 5 males; age range 19-40 years; mean age $M=25.21$, $SD=5.19$) at the Department of Psychology, Stockholm University, participated in Experiment 3 for course credit.

**Materials**
The olfactory materials used in Experiment 3 consisted of the following five concentrations of whiskey: 4%, 8%, 16%, 32%, and 64% whiskey. All odors were kept in nontranslucent glass jars to prevent visual inspection. Each jar contained 10 ml of whiskey. Tap water was used as solvent. As visual stimuli, a black picture in which the
transparency was changed to 96%, 92%, 84%, 68% and 36% was used. In Experiment 3, one tone in 400Hz was created using the program NCH Tone Generator. This tone was presented in different volume: 0dB, -6dB, -12dB, -18dB and -24dB in relation to baseline.

**Results and Discussion**

Figure 4 show the significant and non-significant correlations between stimulus-categories. The same analyses as in Experiment 1 were used. The results showed a negative significant correlation between olfaction and audition ($r = -0.529, t_{13} = -5.651, p < .001$), a high concentration of an odor corresponded to a high volume sound and a low concentration of an odor corresponded to a low volume sound, and a negative significant correlation between audition and olfaction ($r = -0.366, t_{13} = -2.310, p < .05$), a high volume sound corresponded to a high concentration of an odor, and a low volume sound corresponded to a low concentration of an odor. All other correlations were non-significant. Here the results show significance in both ways between audition and olfaction, and compared to Experiment 1 and 2, the significant results between vision and olfaction now disappear altogether. Since the auditory stimuli were changed to volume instead of frequency for Experiment 3 and 4, this is the most likely explanation as to why this occurred. It is also important to note that the correlation between vision and audition disappears. This could be due to there not existing any crossmodal correspondences between color intensity and sound volume, or, that subjects judge these correspondences subjective: for some, a light-grey patch goes together with a loud sound, for others, a light-grey patch goes together with a low sound. Furthermore, the correlations between olfaction and audition speak in favor of existing correspondences between these two modalities regarding this particular dimension.

![Figure 4. Correlations between audition, olfaction and vision, Experiment 3.](image-url)
Experiment 4
The aim of Experiment 4 was to replicate Experiment 3, and to investigate potential influences of olfactory quality on crossmodal correspondence. For this purpose the olfactory stimuli was changed from whiskey to n-butanol.

Method

Participants
Fourteen students (10 females and 4 males; age range 20-37 years; M=26.46, SD=4.94) at the Department of Psychology, Stockholm University, participated in Experiment 4 for course credit.

Materials
In Experiment 4, the olfactory stimuli consisted of the following concentrations of n-butanol (i.e., custom specified and factory calibrated Sniffin’ Sticks): 4%, 8%, 16%, 32%, and 64%. All else remained the same as in Experiment 3.

Results and Discussion
Figure 5 show the significant and non-significant correlations between stimulus-categories. The same analyses as in Experiment 1 were used. The results show a positive significant correlation between olfaction and vision \( (r = 0.329, t_{13} = 2.700, p < .05) \), a high concentration of an odor corresponded to a dark grey picture and a low concentration of an odor corresponded to a light grey picture, and a negative significant correlation between audition and olfaction \( (r = -0.231, t_{13} = -2.230, p < .05) \), a high volume sound corresponded to a high concentration of an odor and a low volume sound corresponded to a low concentration of an odor. All other correlations were non-significant. The difference between Experiment 3 and 4 is that the odor was changed from whiskey in Experiment 3 to Sniffin’ Sticks in Experiment 4. The results largely replicate those of Experiment 3. Although, a new significant correlation can now be seen, that between olfaction and vision; which may be due to chance considered the earlier results in Experiment 1-3. Also, the correlation between olfaction and audition disappears, maybe due to n-butanol being an overall weaker odor than whiskey.
**General Discussion**

The purpose of the present study was to investigate potential crossmodal correspondences between audition, vision, and olfaction, regarding different concentration of stimuli (i.e., light – dark grey color patches, low pitched/volume sounds – high pitched/volume sounds, high concentrations of an odor – low concentrations of an odor). The hypotheses were that high concentrations of odors, low pitch/high volume and dark color-patches, and vice versa, would correspond to each other. Although all of the hypotheses did not receive support, the results show a distinct pattern such that: in Experiment 1 and 2 significant correlations could only be found between vision and audition. In Experiment 3 and 4, where instead the auditory stimuli consisted of a sound in different volume, these results disappear, and instead significant correlations between olfaction and audition/vision are seen.

What may cause these selective correspondences? There could be many possible reasons for this; following are a few of the more likely explanations: (i) that there exists corresponding dimensions between sensory modalities, although not between auditory frequency/grey pictures and olfaction, at least not for whiskey or n-butanol. This would contradict an earlier study by Belkin et al. (1997), who found that people assign certain frequencies to different odors. Although, the designs of the experiments differ, and so does the presented olfactory stimuli; thus, no conclusive inferences can be made. Moreover, the significant correspondences between auditory volume and olfaction (Experiment 3 and 4) add some validity to hypothesis of corresponding dimensions (e.g., Spence, 2011), although along different dimensions; for example, elevation, intensity, or size. It may just be that the sensory information presented in Experiment 1 and 2 does not map onto the same dimensions. (ii) It could be a problem of discriminability, that the odor concentrations here used were too hard to tell apart,
which was mentioned by some participants on debriefing after the experiment. In particular, this was the case for the Sniffin’ Sticks, which was perceived to be weaker than the whiskey even though the same concentrations as for the whiskey were used. This can also be seen on the results: the correlations are generally stronger for olfaction – vision/audition in Experiment 1 and 3, where whiskey was used, compared to Experiment 2 and 4. This could also be because of the qualitatively different odors presented (e.g., whiskey and n-butanol). Interestingly, the significant correlations between olfaction/audition in Experiment 3 and 4, has, to our knowledge, never before been examined. According to the results, there seems to exist dimensions of sound volume, and dimensions of odor-strength, that corresponds to each other. (iii) Furthermore, the non-significant results between olfaction and vision in all experiments except one may be because no congruency effects exists between black/grey color and the particular odors used here; using a forced-choice method then made the participants choose in random. Gilbert et al. (1996) showed robust associations between 20 different odors and colors; more recently, Demattè et al. (2006b) used the method of a speeded classification task, finding that congruent pairs of visual and olfactory stimuli (e.g., the color yellow and a smell of lemon) was faster identified than pairs that were incongruent (e.g., the color red and a smell of lemon). Viewed in light of this, the results presented here could be due to non-existing congruency effects between stimuli used in these particular experiments. Following Spence (2011) suggestion that there exists three types of crossmodal correspondences, structural, statistical and semantic, this could indicate that no statistical (naturally occurring) congruency effects exists between these particular stimuli. And lastly, (iv), that the method and design used here was not adequate to test these particular hypotheses.

Regarding the non-significant results between vision/audition in Experiment 3 and 4, the same explanation used above may be applied here, although these results are also somewhat contradictory to the literature. Stevens & Marks (1965) used a paradigm where the participants matched the brightness of a light to a volume of a sound. The differences in the design of the experiments could account for the results presented here; it could also be that, as speculated above, no corresponding dimensions exists between sound volume and the particular color-patches used here. One possible weakness of the design used for this experiment is the forced-choice procedure. If no actual correspondences exist, the subjects still have to make a choice, which due to an expectancy bias may have influenced the results. If there were actual correspondences between, say, sound frequency and grayscale, the participants may have used the same strategy for the remaining matching’s, because of a belief that this was what was expected from them. One alternative procedure could have been the method used by Belkin et al. (1997), i.e., letting the subjects choose what frequency actually corresponded to each odor. Those frequencies could later on be used in the experimental design presented here. In addition, to further test the hypotheses stated here, one could use larger steps of concentrations of odors (e.g., 10%, 50%, and 90%). The odors would then be easier discriminated, giving more conclusive results. Further research is needed to validate the results, and also to see whether the olfactory stimuli used here can be associated to different frequencies. Moreover, replications of the present study but with different odors would be of help in explaining the results.

In conclusion, the present study replicated earlier results using a novel method, testing
all correspondences in the same trial and varying the concentration for all stimuli. The study also demonstrated that there exist correspondences between vision and frequency of a sound (Experiment 1 and 2). Furthermore, correspondences between the concentration of an odor and the volume of a sound were found (Experiment 3 and 4). Also, one significant correlation between olfaction and vision (Experiment 4) was found, although more research is needed to confirm this result. It is suggested that crossmodal correspondences are observed when sensory information from two or more modalities map onto the same dimension.

References


Zampini, M., & Spence, C. (2010). Assessing the Role of Sound in the Perception of Food and