



Nonlinear response speedup in bimodal visual-olfactory object identification

Richard Höchenberger^{1*}, Niko A. Busch^{2,3} and Kathrin Ohla¹

¹ Psychophysiology of Food Perception, German Institute of Human Nutrition Potsdam-Rehbrücke (DIFE), Nuthetal, Germany, ² Institute of Medical Psychology, Charité - Universitätsmedizin Berlin, Berlin, Germany, ³ Berlin School of Mind and Brain, Humboldt University, Berlin, Germany

OPEN ACCESS

Edited by:

Andriy Myachykov,
Northumbria University, UK

Reviewed by:

Wen Li,
Florida State University, USA
Ashley James Chapman,
Northumbria University, UK

*Correspondence:

Richard Höchenberger,
German Institute of Human Nutrition
Potsdam-Rehbrücke (DIFE),
Arthur-Scheunert-Allee 114–116,
14558 Nuthetal, Germany
hoechenberger@dife.de

Specialty section:

This article was submitted to
Cognition,
a section of the journal
Frontiers in Psychology

Received: 10 April 2015

Accepted: 14 September 2015

Published: 30 September 2015

Citation:

Höchenberger R, Busch NA and
Ohla K (2015) Nonlinear response
speedup in bimodal visual-olfactory
object identification.
Front. Psychol. 6:1477.
doi: 10.3389/fpsyg.2015.01477

Multisensory processes are vital in the perception of our environment. In the evaluation of foodstuff, redundant sensory inputs not only assist the identification of edible and nutritious substances, but also help avoiding the ingestion of possibly hazardous substances. While it is known that the non-chemical senses interact already at early processing levels, it remains unclear whether the visual and olfactory senses exhibit comparable interaction effects. To address this question, we tested whether the perception of congruent bimodal visual-olfactory objects is facilitated compared to unimodal stimulation. We measured response times (RT) and accuracy during speeded object identification. The onset of the visual and olfactory constituents in bimodal trials was physically aligned in the first and perceptually aligned in the second experiment. We tested whether the data favored coactivation or parallel processing consistent with race models. A redundant-signals effect was observed for perceptually aligned redundant stimuli only, i.e., bimodal stimuli were identified faster than either of the unimodal components. Analysis of the RT distributions and accuracy data revealed that these observations could be explained by a race model. More specifically, visual and olfactory channels appeared to be operating in a parallel, positively dependent manner. While these results suggest the absence of early sensory interactions, future studies are needed to substantiate this interpretation.

Keywords: multisensory integration, olfaction, visual-olfactory, race model, response time

1. Introduction

Olfactory and visual sensory information are continuously flooding the brain and are, therefore, often experienced with a marked temporal overlap or even simultaneously. Both the smell and visual appearance serve a vital function in the localization of food, the assessment of edibility, as well as the identification of potential environmental hazards, thereby allowing for fast and appropriate behavior not only limited to food-choice. The integration of redundant sensory information by the neural system has been proven beneficial for perception and subsequent behavior: it speeds up processing and improves accuracy. However, it is unclear whether this holds true for the combination of olfaction and vision.

Recent studies have shown that odors modulate visual perception and performance, particularly by directing attention to and influencing the saliency of a congruent visual object, e.g., during attentional blink (Robinson et al., 2013), binocular rivalry (Zhou et al., 2010, 2012), spatial

attention and visual search (Chen et al., 2013), and eye movements (Seo et al., 2010). These effects occur even when odors are task-irrelevant and suggest spontaneous binding between visual and olfactory inputs (Zhou et al., 2012). In contrast, odor perception is not only influenced by vision (and the other senses), but odor identification also critically depends on additional information because odors in isolation are notoriously ambiguous (Cain, 1979). Observations that humans have difficulties identifying and discriminating odors in the absence of additional information (Davis, 1981) and that color cues (Zellner et al., 1991), verbal labels (Herz and von Clef, 2001) and images (Gottfried and Dolan, 2003) assist odor perception corroborate this notion. Most previous studies investigated modulatory effects of visual cues on olfaction and their interaction at cognitive levels, when semantic representations were available. It remains unknown whether sensory information from the olfactory and visual modalities is in fact pooled at early perceptual stages, that is, integrated.

1.1. How can we Investigate Whether Multisensory Integration is Taking Place?

Multisensory integration has been mostly studied between the non-chemosensory modalities vision, hearing, and somatosensation; these senses have been shown to interact already at the level of the superior colliculi (Stein and Meredith, 1990). Classically, super-additive responses, i.e., more than the sum of the parts, are considered an indication of multisensory integration. Key aspects governing multisensory integration are the so-called principles of spatial and temporal proximity: stimuli presented at the same location and at the same time, respectively, most likely belong to the same object and are therefore more likely to be bound together to a unitary percept (Stein and Meredith, 1993). Additionally, Meredith and Stein (1986) found that cells in the superior colliculus produced the strongest response amplification for the weakest stimuli, a principal phenomenon called *inverse effectiveness*. While these observations could be replicated on the behavioral level in numerous studies, it has been suggested that these findings might largely be statistical artifacts (Holmes, 2007). While imaging studies have mostly focused on superadditive effects when trying to identify functional correlates of multisensory integration, it is unclear whether the results from single-neuron recordings can be readily transferred to the cortical level (Laurienti et al., 2005) and behavior.

1.2. Response Facilitation Can Serve as a Possible Measure of Multisensory Integration

Stimulus detection, on average, is faster and more accurate in situations where the target is presented redundantly, i.e., on several sensory channels. In the multisensory literature, this facilitation is commonly called *redundant-targets effect* (RTE) or *redundant-signals effect* (RSE); both terms are largely used synonymously. In the remainder of this paper, we will refer to these effects of multisensory processing exclusively as redundant-signals effects. The response speedup is commonly explained by assuming that an internal decision criterion is reached faster when multiple targets are presented simultaneously, compared

to the single-target situation. Similarly, redundant information reduces stimulus ambiguity, hence allowing for a higher accuracy of responses.

However, RSEs can result from *statistical facilitation* merely due to probability summation alone. A popular probability summation model was introduced by Raab (1962) with the idea of a *race* between parallel single-target detection processes during a multiple-target situation. The process finishing first “wins the race,” elicits a response, and, therefore, determines the behavioral response time. These so-called race models operate according to a *separate-activation model* with a *first-terminating stopping rule* (see e.g., Colonius and Vorberg, 1994). They implicitly assume unlimited-capacity processing (Colonius, 1990), meaning that the speed of one detection process is not influenced by other, simultaneous, detection processes. For example, detection of a unimodal target should happen at the same speed as detection of the same target in a multimodal situation. Therefore, if RT distributions of the single-target detection processes overlap, the observed RTs in redundant-target trials will, on average, be faster than the unimodal RTs. “Slow” responses of one single-target detection process can be replaced by “faster” responses of another, simultaneous detection process. The observed RT speedup would thus be a statistical artifact only. In sum, an RSE that can be fully accounted for by a race model does not provide strong evidence for multisensory integration.

Nevertheless, integration can be inferred if RTs are faster than predicted by race models. Specifically, Miller (1982) derived an upper bound to the bimodal RT speedup possible in *any* race model, the so-called *race model inequality* (RMI) or *Miller bound*. It is based on the assumption of maximum negative dependence between the channel processing speeds (Colonius, 1990), that is, if the participant detects a signal on one channel at a given bimodal trial, the other channel will fail to detect the target. Violations of this criterion, i.e., faster responses than predicted by the RMI, support *coactivation models*. They demand that processing of different sensory channels be pooled prior to the decision stage and therefore refute all race models in favor of “true” multisensory integration. Satisfaction of the RMI, on the contrary, does not necessarily exclude coactive processing.

Numerous studies have investigated response facilitation to bimodal stimuli in the visual, auditory, and somatosensory modalities (see e.g., Gielen et al., 1983; Miller, 1986; Forster et al., 2002; Diederich and Colonius, 2004). Whether the combined presentation of congruent (that is, redundant) visual-olfactory information can likewise facilitate object perception remains unclear and was investigated with the present study. Specifically, we tested the hypothesis that bimodal visual-olfactory object identification is facilitated compared to identification of either of the unimodal constituents alone.

Furthermore, we examined whether facilitation is more pronounced for perceptually aligned compared to physically aligned stimuli. For this, we conducted two experiments in which the bimodal constituents were either presented physically (Experiment 1) or perceptually (Experiment 2) simultaneously. We compared the observed RTs and response accuracies to the predictions of different models of probability summation.

2. Results

Seven participants smelled and viewed different food objects presented either alone as unimodal visual (V) or olfactory (O) stimuli or as congruent bimodal combinations (OV) and performed a speeded two-alternative forced-choice (2-AFC) object identification task. Stimulus strength was adjusted to achieve approximately 75% accuracy. The biggest RSE for response time (RT) can be observed when the RT distributions of the unimodal constituents overlap largely (Hilgard, 1933; Hershenson, 1962; Raab, 1962; Miller, 1986; Colonius, 1990; Diederich and Colonius, 2004; Gondan, 2009). Therefore, we conducted two separate experiments, in which OV stimuli consisted of *physically* (Experiment 1; **Figure 1**, left) or *perceptually* (Experiment 2; **Figure 1**, right) aligned unimodal constituents. Perceptual alignment was achieved by introduction of a stimulus-onset asynchrony (SOA) equal to the RT differences between the unimodal stimuli.

For perceptually aligned OV stimuli (Experiment 2) we observed a significant RSE for RTs [RSE = 54 ms, $t_{(6)} = 3.05$, $p = 0.02$], that is a speedup of median RTs to bimodal OV compared to the fastest unimodal stimuli (**Figure 2A**). No RSE was found for physically aligned OV stimuli during Experiment 1 [RSE = 0 ms, $t_{(6)} = 0.02$, $p = 0.98$]. Accuracy showed no RSE in either Experiment [Experiment 1: RSE = 1.8 %-points, $t_{(6)} = 0.73$, $p = 0.49$; Experiment 2: RSE = -0.6 %-points, $t_{(6)} = -0.29$, $p = 0.78$; **Figure 2B**].

Experiment 1 yielded a significant difference between response times (RT) to unimodal V and O stimuli, indicating only little overlap between the unimodal RT distributions; V stimuli were perceived 362 ms faster than O stimuli [$t_{(6)} = -4.72$, $p < 0.01$; **Figure 3A**]. By contrast, we could observe a markedly reduced difference between the SOA-corrected unimodal RTs in Experiment 2 of only 74 ms [$t_{(6)} = -2.48$, $p < 0.05$; **Figure 3B**], indicating a strong overlap of unimodal RT distributions.

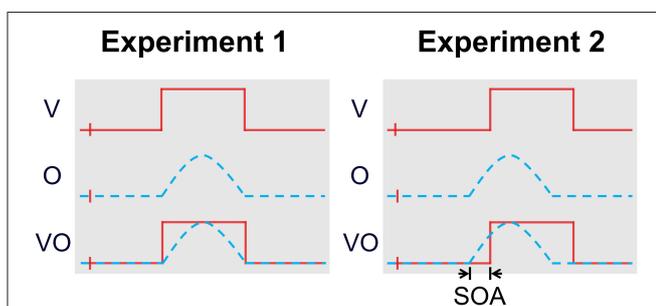


FIGURE 1 | Stimulus timing in unimodal and bimodal trials of Experiments 1 and 2 (schematic). The start of a trial is marked with a small vertical line and identifies the onset of the fixation cross. The visual and olfactory constituents were presented physically simultaneously in the bimodal trials in Experiment 1, and perceptually simultaneously in Experiment 2. The SOAs employed in Experiment 2 were individually estimated for every participant and object (banana and lemon, respectively) during Experiment 1. Note that the depicted delayed presentation of the visual stimulus in Experiment 2 caused the fixation cross to be displayed for a longer duration before stimulus onset in the unimodal vision-only condition in order to ensure context-invariance.

These performance differences were clearly reflected in the cumulative RT distributions: While the bimodal distribution mostly followed the visual distribution in Experiment 1, it was shifted toward faster responses throughout its whole range in Experiment 2 (**Figures 3C,D**).

We next tested whether the bimodal RT speedup could be explained by statistical facilitation in a separate-activation model with unlimited capacity (race model). The theoretical upper performance limit was given by the *Miller bound*: If observed responses were *faster* than this boundary at any time, all race models could be ruled out at once, and the system would be assumed to be super-capacity at this time (Townsend and Nozawa, 1995; Townsend and Wenger, 2004). Additionally, we compared our data to a lower performance bound proposed by Grice et al. (1984b), referred to as *Grice bound*. It assumes that responses in the bimodal situation should be at least as fast as responses to the fastest unimodal constituents. If responses were *slower* than this boundary, the system would be assumed to be limited-capacity (Townsend and Wenger, 2004) at this time. Both super-capacity and limited-capacity processing violate the assumption of context-invariance, invalidating an essential requirement of race models (Colonius, 1990). We found that the observed bimodal RTs did not exceed the *Miller* or the *Grice bounds* significantly in both experiments, indicating parallel processing of the visual and olfactory channels (**Figures 3E,F**, and **Table 1**). Our data could thus be attributed for by a race model.

Colonius (1990) pointed out that the *Miller* and *Grice bounds* can only be reached under the implicit assumption of perfect negative and positive, respectively, dependence between channel processing speeds. To gain further insight into the

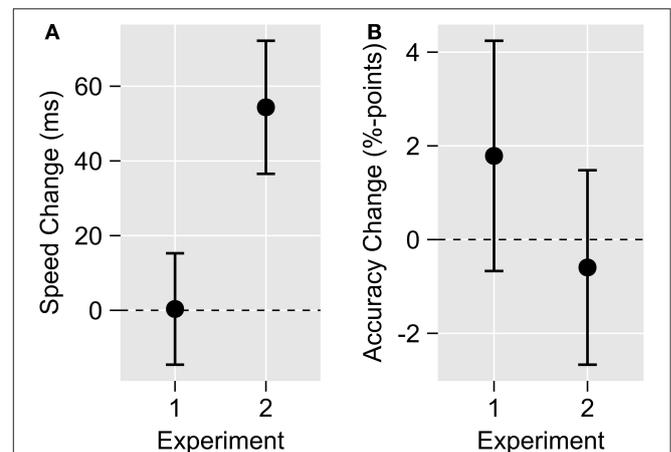


FIGURE 2 | Mean redundant-signals effects in Experiments 1 and 2. Positive values indicate a bimodal facilitation, negative values a bimodal impairment relative to the unimodal constituents. **(A)** No redundancy gain in response speed was observed for physically simultaneous bimodal stimulation (Experiment 1), but it was clearly evident for perceptually simultaneous stimulation (Experiment 2). **(B)** For both physically simultaneous (Experiment 1) and perceptually simultaneous bimodal stimulation (Experiment 2), no significant accuracy improvement could be observed. Data were calculated individually for each participant and object, and subsequently averaged. Error bars show standard error of the mean.

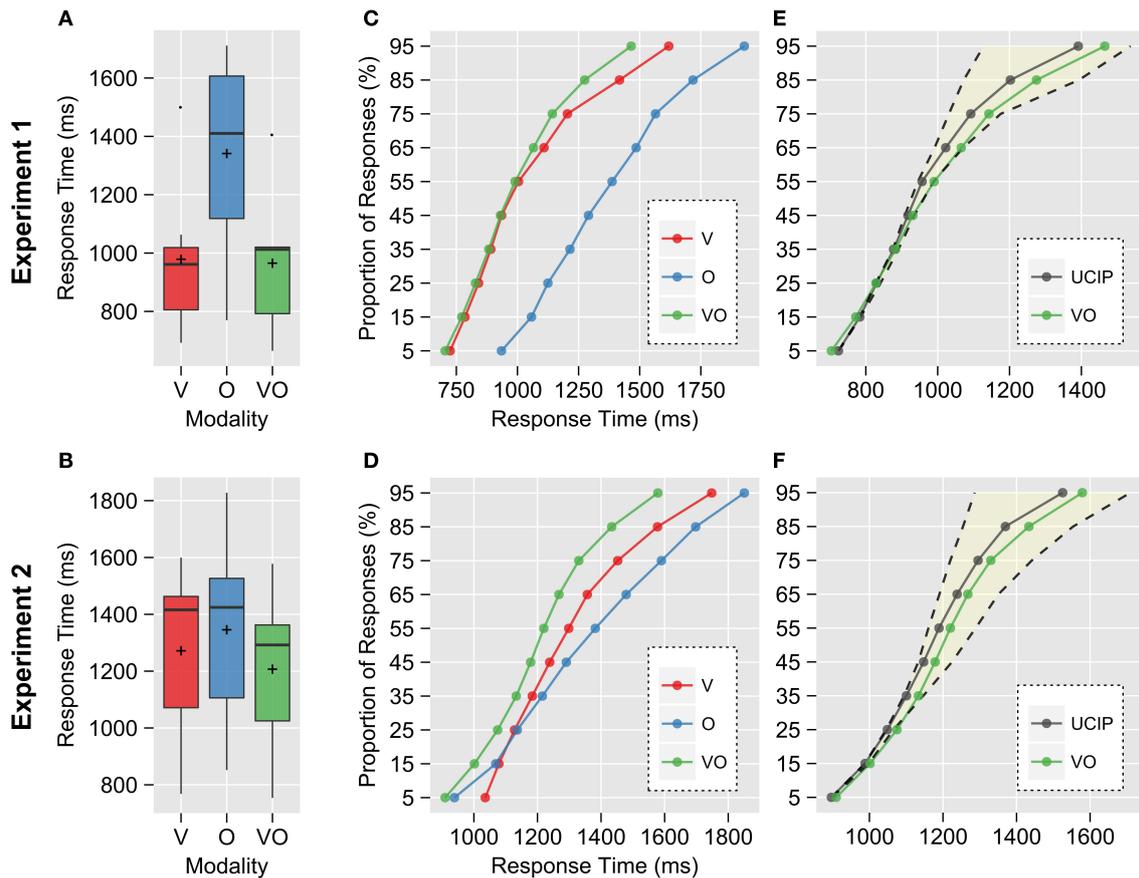


FIGURE 3 | Comparison of observed response times (RTs) in the visual (V), olfactory (O), and bimodal (VO) trials. Visual RTs were corrected for SOA. **(A)** RT distributions for V and VO appeared to be almost identical, with the O distribution shifted to much slower RTs and very little overlap with V, indicating that responses in VO trials were mostly driven by the visual constituent. The crosses depict group means. **(B)** Individual timing adjustments by introduction of an SOA aligned the unimodal distributions, suggesting perceptual simultaneity. **(C,D)** Empirical cumulative RT distributions. Quantile values were averaged across participants. In Experiment 1, the VO distribution seemed to follow V up to the 55% quantile. Approximately at the same time, the fastest olfactory responses could be observed, i.e., the unimodal constituents were starting to perceptually overlap. Coincidentally, the VO distribution started to diverge from V, and shifted to faster responses. This bimodal speedup is not reflected in the global (mean) RSE. In Experiment 2, the VO distribution was shifted to faster responses relative to both unimodal distributions across its whole range. **(E,F)** Comparison of the unlimited capacity, independent, parallel (UCIP) model prediction with the observed bimodal data. The highlighted area depicts the possible phase space under the assumption of separate-activation models with unlimited capacity and a first-terminating time rule, but possibly dependent processing (i.e., possible race models would have to lie within this area); accordingly, the dashed line to the left shows the Miller and the right the Grice bound (upper and lower performance limits, respectively).

underlying processing mechanisms, we compared our data to a model assuming uncorrelated processing between the visual and olfactory channels, the so-called *unlimited-capacity, independent, parallel* (UCIP) model. The bimodal RTs were slower than predicted by this model in the 75 % quantile in Experiment 1 and from the 45 % to the 85 % quantiles in Experiment 2 (all $p < 0.05$). However, only the deviation in the 75 % quantile in Experiment 2 survived Holm-Bonferroni correction for multiple testing ($p_{\text{corr}} = 0.03$). All comparisons are summarized in **Table 1**. Because the deviations of the observed data from the model predictions are shifted in direction of the *Grice bound*, i.e., toward perfect positive dependence, the results suggest a race model with positively correlated channel processing speed between the visual and olfactory channels (see Grice et al., 1984a).

Next, the accuracy data (grand means shown in **Figures 4A,B** for Experiments 1 and 2, respectively) were compared to models of probability summation. We first adopted equivalents of the *Miller* and *Grice bounds* to derive upper and lower performance limits, respectively (Colonius, 2015). The upper bound was at 100 % accuracy in both experiments and therefore never violated; observed accuracies were significantly below this bound [Experiment 1: $\Delta = -13.5\%$ -points, $t_{(6)} = -8.95$, $p < 0.001$; Experiment 2: $\Delta = -13.9\%$ -points, $t_{(6)} = -6.76$, $p < 0.001$]. The lower bound was never significantly violated [Experiment 1: $\Delta = 1.8\%$ -points, $t_{(6)} = 0.73$, $p = 0.49$; Experiment 2: $\Delta = -0.6\%$ -points, $t_{(6)} = -0.29$, $p = 0.78$]. Note that the lower bound was identical to the baseline used earlier to identify an RSE for accuracy. These results suggest that probability summation could in fact explain the observed bimodal accuracies.

TABLE 1 | Each participant contributed one quantile value, following the procedure from Ulrich et al. (2007).

| Experiment | Quantile (%) | VO–Miller bound | | | VO–Grice bound | | | VO–UCIP | | |
|------------|--------------|-----------------|-------|-------|----------------|-------|-------|---------------|-------|---------------|
| | | Δ (ms) | t | p | Δ (ms) | t | p | Δ (ms) | t | p |
| 1 | 5 | –20 | –1.68 | 0.144 | –20 | –1.68 | 0.145 | –20 | –1.68 | 0.144 |
| | 15 | –11 | –1.15 | 0.295 | –13 | –1.52 | 0.180 | –11 | –1.18 | 0.281 |
| | 25 | 0 | 0.04 | 0.971 | –13 | –1.99 | 0.094 | –2 | –0.27 | 0.800 |
| | 35 | 8 | 0.82 | 0.443 | –8 | –1.43 | 0.204 | 5 | 0.58 | 0.583 |
| | 45 | 22 | 1.30 | 0.242 | –4 | –0.36 | 0.732 | 14 | 0.92 | 0.391 |
| | 55 | 45 | 2.30 | 0.061 | 2 | 0.11 | 0.915 | 33 | 1.85 | 0.114 |
| | 65 | 74 | 2.65 | 0.038 | –6 | –0.21 | 0.844 | 43 | 1.79 | 0.123 |
| | 75 | 111 | 4.26 | 0.005 | –33 | –1.05 | 0.333 | 50 | 2.62 | 0.039 |
| | 85 | 203 | 4.11 | 0.006 | –118 | –1.84 | 0.116 | 73 | 1.77 | 0.128 |
| 95 | 337 | 4.41 | 0.004 | –72 | –0.91 | 0.400 | 74 | 0.94 | 0.384 | |
| 2 | 5 | 13 | 0.77 | 0.472 | 11 | 0.65 | 0.543 | 13 | 0.77 | 0.473 |
| | 15 | 14 | 0.92 | 0.395 | 4 | 0.24 | 0.822 | 13 | 0.86 | 0.422 |
| | 25 | 29 | 2.01 | 0.091 | 9 | 0.52 | 0.622 | 26 | 1.71 | 0.138 |
| | 35 | 39 | 2.20 | 0.070 | –13 | –1.26 | 0.255 | 33 | 1.99 | 0.093 |
| | 45 | 46 | 3.98 | 0.007 | –46 | –3.05 | 0.022 | 31 | 3.49 | 0.013 |
| | 55 | 60 | 4.08 | 0.007 | –69 | –2.54 | 0.044 | 31 | 3.01 | 0.024 |
| | 65 | 77 | 5.08 | 0.002 | –83 | –2.92 | 0.027 | 29 | 2.92 | 0.027 |
| | 75 | 110 | 5.92 | 0.001 | –112 | –3.39 | 0.015 | 35 | 4.82 | *0.003 |
| | 85 | 179 | 6.16 | 0.001 | –120 | –2.39 | 0.054 | 64 | 3.47 | 0.013 |
| 95 | 293 | 4.89 | 0.003 | –131 | –2.09 | 0.082 | 54 | 1.37 | 0.218 | |

Negative t -values for the Miller and positive t -values for Grice bound comparisons indicate violations of the respective bounds. Significant model violations in bold; the asterisk marks a significant violation after Holm-Bonferroni correction for multiple testing.

The data were then compared to a model predicting stochastic independence, equivalently to the UCIP model employed for RTs (Stevenson et al., 2014). Bimodal accuracy was lower than predicted by the model in both Experiment 1 [$\Delta = -9.0\%$ -points, $t_{(6)} = -5.97$, $p < 0.001$; **Figure 4C**] and Experiment 2 [$\Delta = -10.6\%$ -points, $t_{(6)} = -4.99$, $p < 0.01$; **Figure 4D**]. In line with the RT data, the accuracy data indicate that the visual and olfactory channels are stochastically dependent.

3. Discussion

The present study found a bimodal response facilitation for perceptually, but not for physically aligned bimodal visual-olfactory stimuli. The facilitation could be accounted for by race models assuming probability summation across positively dependent processing channels. Thus, the results yielded no proof of coactivation.

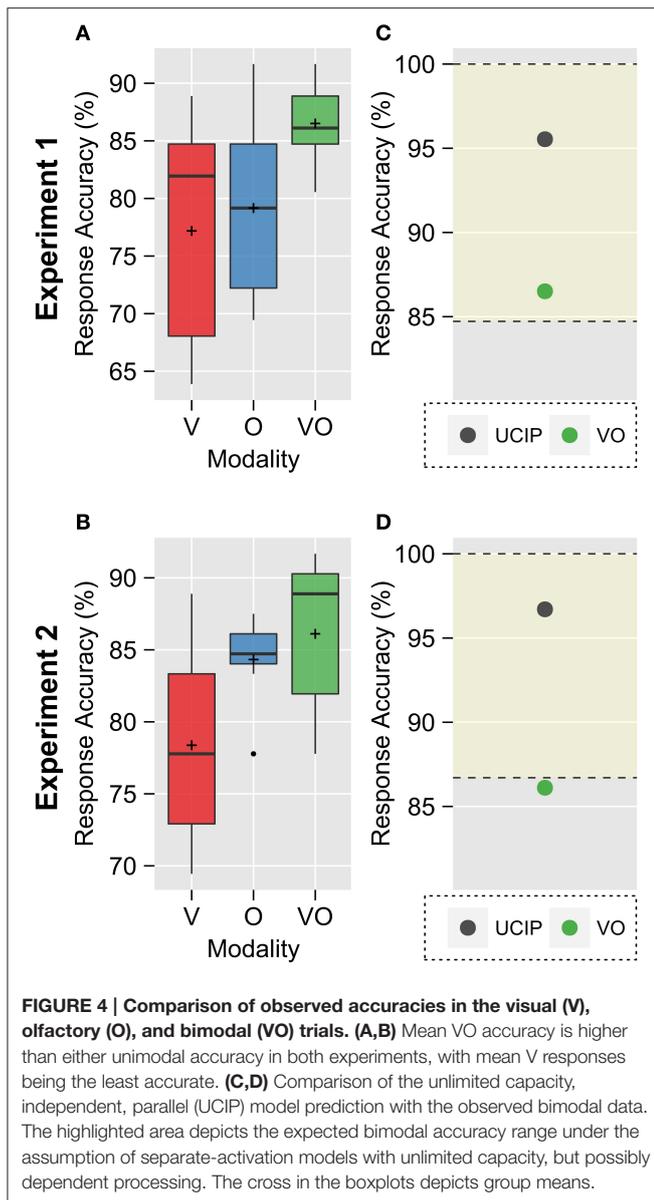
The observation of a significant bimodal visual-olfactory response speedup indicated by an RSE for perceptually, but not for physically aligned unimodal constituents suggests that temporal proximity subserves visual-olfactory response facilitation. Increasing temporal parity amplifies multisensory interactions in other sensory modalities, e.g., for visual-tactile (Forster et al., 2002) or visual-auditory (Lovelace et al., 2003) stimuli albeit the temporal binding window, i.e., the range of inter-stimulus intervals over which multisensory stimuli are

integrated, is not universal. While multisensory binding windows as large as several hundred milliseconds exist for example in audio-visual speech perception (see e.g., van Wassenhove et al., 2007), the effects of stimulus timing on visual-olfactory perception are unknown. To our knowledge, this is the first study demonstrating that perceptual, rather than physical, simultaneity is vital to elicit an RSE for bimodal visual-olfactory objects.

However, the response speedup could be the result of statistical facilitation alone and is not necessarily proof of neural integration processes. Therefore, we examined whether the present data could be explained by race models, or if we could find evidence for coactivation.

Response time distributions never significantly exceeded the *Miller bound*. We can therefore exclude coactivation as a possible explanation of the observed RSEs. Further we can exclude strictly limited-capacity processing over an extended period of time because the *Grice bound* was never violated (Townsend and Wenger, 2004). Taken together, the observed bimodal response times are consistent with separate-activation models with a first-terminating time rule and unlimited-capacity processing, i.e., race models (Miller, 1982; Grice et al., 1984a; Colonius, 1990).

Classically, it has been shown that violations of the *Miller bound* are more easily produced in go/no-go tasks due to the absence of “response competition” (Grice and Canham, 1990; Grice and Reed, 1992). Yet, race models can successfully be rejected in choice response time studies as well (see e.g., Miller, 1982; Hecht et al., 2008; Girard et al., 2011).



No change in response accuracy was observed in the bimodal conditions, compared to unimodal stimulation. This finding is in contrast to previous reports of improved accuracy for multisensory stimuli. A possible reason for this discrepancy might be that olfaction and vision do not integrate in the same way as other senses. However, it is also possible that we were not able to observe improved accuracy simply for statistical reasons due to the low number of trials (owing to the long inter-trial intervals, ITIs, necessary for olfactory stimuli) and high inter-subject variability.

Comparison of the observed bimodal response time distributions to a more restrictive race model assuming stochastic independence of channel processing speed (UCIP model) revealed significantly slower responses than predicted in both experiments, suggesting positively dependent channel

processing speeds between the visual and olfactory channels (Grice et al., 1984a). Although only the deviation in one quantile in Experiment 2 was significant after correction for multiple testing, the additional finding of lower bimodal response accuracies than predicted further corroborates the assumption of a possibly positive stochastic dependence of visual and olfactory processing.

In contrast, the bimodal combination of odor and taste stimuli yielded faster responses than predicted by a UCIP model in a recent study (Veldhuizen et al., 2010). Notably, odor and taste perception are closely intertwined; evidence exists for direct and indirect anatomical connections between the primary gustatory and olfactory cortices (Rolls and Baylis, 1994; Shepherd, 2006) as well as for convergence areas responding to both smell and taste, for example in the orbitofrontal cortex (OFC) (O'Doherty et al., 2001; de Araujo et al., 2003; Small and Green, 2012), the anterior insula, and frontal and parietal opercula (Small et al., 1999; Cerf-Ducastel and Murphy, 2001; Poellinger et al., 2001). Perceptually, the combined odor-taste experience typically exceeds the sum of the two chemosensory modalities, being perceived as more *Gestalt*-like, intense and rewarding, and yields superadditive activation in the frontal operculum (Seubert et al., 2015). Although no monosynaptic connection between the primary visual and olfactory cortices has been found, the perirhinal cortex is a prime candidate as a processing hub between the visual and olfactory modalities due to its numerous reciprocal connections, particularly with the inferior temporal cortex. The inferior temporal cortex is involved in object perception (Grill-Spector and Weiner, 2014) and associations of sensory representations, and a subdivision, the rhinal cortex, has been proven critical for the association of flavor with visual food objects in monkeys (Parker and Gaffan, 1998).

3.1. Conclusion

The present data are consistent with models of parallel processing with unlimited-capacity and positive dependence between the visual and olfactory channels. Notably, these models do not refute the possibility of coactive processing. Although odor perception is highly ambiguous and susceptible to other sensory information (Herz and von Clef, 2001), the olfactory stimuli may in fact have contributed to the bimodal object identification by generating further perceptual evidence, allowing an internal decision criterion to be reached faster. This assumption is supported by the observation of positive channel dependence, indicating that the identification of the visual and olfactory constituents in bimodal trials co-occurs. The objects used in the present study carried a semantic meaning, which had to be decoded before mapping it to the appropriate response button. Semantic representations emerge only at later stages in the perception process (Olofsson, 2014). Further, no direct connections between the visual and olfactory cortices have been discovered yet, questioning the plausibility of early bimodal visual-olfactory interactions. Future studies will have to show whether the present findings are transferable to other stimulus objects, SOAs, and experimental tasks.

4. Materials and Methods

4.1. Participants

Eight participants completed the study; one participant was excluded because his accuracy was far below chance level for the unimodal olfactory lemon stimulus in both experiments (mean accuracy was approx. 33%); data of seven participants (4 female; age in years: 29.9 ± 2.4 SD, range: 26–32; all right-handed) are reported here. Participants were recruited from the German Institute of Human Nutrition and local universities; they gave written informed consent and received compensatory payment. They reported no neurological disorders or chronic diseases, in particular no smell impairment, and normal or corrected-to-normal vision. The study was conducted in accordance with the requirements of the revised Declaration of Helsinki and had been approved by the ethics committee of the German Society for Psychology (DPGs).

4.2. Stimuli

4.2.1. Visual Stimuli

Six images (three different images of bananas and lemons, respectively) with different complexities were selected from the Food-pics database (no. 276, 282, 341, 379, and 415; Blechert et al., 2014) or purchased online. Images displayed a food object centered on a white background. They were resized to 1024×1024 pixels and converted to grayscale. A Gaussian blur (order 0, $\sigma = 3$) was applied to remove sharp edges. The fast Fourier transform (FFT) of all images was calculated and the phase space was randomly scrambled. The inverse FFT of the image with the scrambled phase yielded blurry images of the food objects with superimposed cloud-like noise patterns. Noise-only images were also derived for every object using the same method, yielding 2×3 target and 2×3 noise-only stimuli in total. The spatial frequency of those noise patterns was similar to the spatial frequency of the original object. Images were presented on a TFT monitor with a resolution of 1680×1050 pixels. The refresh rate was set to 60 Hz. Participants viewed the images at an eye distance of approx. 60 cm, corresponding to an object size of approximately 12° of visual angle, embedded in visual noise of approximately 27° of visual angle.

4.2.2. Olfactory Stimuli

Odorants were 10 mL aliquots of isoamyl acetate (banana; Sigma-Aldrich Chemie GmbH, Steinheim, Germany, CAS 123-92-2) and lemon oil (lemon; same vendor, CAS 8008-56-8) diluted with mineral oil (Acros Organics, Geel, Belgium, CAS 8042-47-5) to produce solutions of 0.1% v/v concentration. The solvent, pure mineral oil, served as neutral control. The odors were congruent to the visual objects banana and lemon; odor intensity was chosen to yield identifiable, yet weak stimuli based on a pilot study ($n = 7$). Odorants were presented birhinally using a custom-built 16-channel air-dilution olfactometer (Lundström et al., 2010). Teflon tubes with an inner diameter of $1/16''$ delivered the odorous air via custom-made anatomically shaped nose pieces into the participants' nostrils. A constant flow of clean air (approximately 0.5 L min^{-1}) was present at all times to rinse the tubing system and the nose. Stimuli were delivered with a flow

rate of approximately 3.0 L min^{-1} , totaling to a flow of about 3.5 L min^{-1} during stimulation. Stimulus timing was measured using a photo-ionization detector (PID; 200B miniPID, Aurora Scientific Inc., Aurora/ON, Canada) and defined as the time point 254 ms after sending the trigger to the olfactometer. To ensure a constant odor concentration and to reduce depletion of head space in the odor jars in the course of the experiment, one of three identical odor jars was used in sequential order from trial to trial.

4.3. Procedure

Participants completed two experimental sessions on separate days. In the first session, a visual identification threshold assessment was conducted, followed by a choice response time (CRT) Experiment in which bimodal stimulus components were presented physically simultaneous. A second CRT experiment with perceptually aligned bimodal stimuli was conducted during the next session. The experiments were carried out in a sound-attenuated experimental booth. Participants were seated centered in front of the screen. Responses were collected using a button box (Serial Response Box, Psychology Software Tools, Sharpsburg/PA, USA) connected to a USB port of the stimulation computer via a serial-to-USB adapter. Timing accuracy was verified to be better than 2 ms. In-ear headphones delivered Brownian noise during the CRT experiments at a volume chosen such that the change in air flow at stimulation on- and offset was inaudible. The stimulation was controlled using PsychoPy 1.79.01 (Peirce, 2009) running on a personal computer.

4.3.1. Visual Threshold Estimation

We adjusted the strength of the noise so that objects could be perceived approximately on every second trial using a QUEST staircase procedure (Watson and Pelli, 1983). The Experiment started with a short practice block, in which all target and noise-only images were presented once. Then, images of objects + noise were presented interleaved with noise-only images (equal proportions) for 900 ms with a randomly varied ITI between 1.5 and 2.0 s during which a white screen was presented. Participants indicated by button press the detection of an object within the noise. The staircase adjusted the strength of the noise to yield a performance level of 50% correct object detection when stimuli were present (false alarms on noise-only trials were very rare, ranging from 0 to approx. 3%, with a grand mean of 1.3%). Separate staircases were run for each of the six different object images. Overall, the threshold procedure entailed 240 trials, 20 repetitions of each of the six images and their respective noise-only images ($2 \times 6 \text{ images} \times 20 \text{ repetitions}$). Participants were allowed a short break; the procedure lasted about 12 m. Note that stimuli yielding 50% accuracy in this detection task are expected to yield approximately 75% performance in the 2-AFC task as used in the main experiment.

4.3.2. Bimodal CRT Experiments

During the CRT experiments, participants were to identify the presented object (banana or lemon) as quickly as possible (while avoiding anticipatory responses) by pressing either of two buttons on the button box. Stimuli were either unimodal visual (V) objects presented at individual 50% identification threshold,

unimodal olfactory (O) objects, or bimodal visual-olfactory (VO) objects. V stimuli were always paired with the neutral control odorant. O stimuli were paired with a randomly assigned noise-only image derived from a visual stimulus of the same object. OV stimuli consisted of the combined presentation of congruent V and O stimuli.

VO stimulus pairs were presented simultaneously in Experiment 1. In Experiment 2, bimodal stimulus timing was adjusted to achieve perceptual simultaneity by introducing an SOA equal to the difference of unimodal median RTs individually for each object and participant. The mean SOAs were 330 ± 295 ms SD for banana, and 395 ± 205 ms SD for lemon. Note that all SOAs were positive, i.e., delaying visual presentation, except for banana in one participant, where the odor had to be presented 182 ms prior to the visual stimulus to achieve perceptual simultaneity. To ensure context-invariance in Experiment 2, we also adjusted the timing of the unimodal stimulus presentations. Specifically, if the estimated SOA indicated a delayed presentation of the visual constituent in bimodal conditions, we also delayed the visual stimulation in the unimodal conditions for the same amount of time (meaning the fixation cross was visible for a longer duration before the stimulus appeared; note that this was also true for the unimodal olfactory stimulation, where the visual stimulus was noise-only). The stimulus timing is illustrated in **Figure 1**.

Each trial started with a fixation cross centered on the screen, which informed participants to prepare and to slowly inhale. At the same time, the air flow through the neutral jar was initiated to remove the tactile cue from the later stimulus presentation. After a random period of 1–2 s, a stimulus (O, V, or VO) was presented for 900 ms. After stimulation, the neutral control odorant was presented for 4.1 s to remove residual odor molecules. The ITI was randomly varied between 20 and 21 s.

The experiments started after a short practice block in which each stimulus combination was presented once. Each Experiment consisted of six blocks during which all stimulus combinations were presented twice and in pseudo-random order, totaling to 216 stimuli (6 blocks \times 2 repetitions \times (6 V + 6 O + 6 VO)), and lasted 95–120 min. Participants were allowed self-paced breaks in the middle of each block and between blocks.

RT measurement started with the onset of the image in V trials and the physical onset of the odorant as determined by PID measurements in O trials. In bimodal trials, RT measurement started with the physical onset of the stimuli (Experiment 1, physically simultaneous presentation), or with the onset of the earlier stimulus (Experiment 2, perceptually simultaneous presentation).

4.4. Data Analysis

Only trials with positive and correct identification responses were analyzed. RT medians and standard deviations (SDs) of the aggregated data were calculated for each of the six conditions (O, V, VO for banana and lemon objects, respectively). All trials with a reaction time deviating more than two SDs from the median were discarded as outliers. In Experiment 1 and 2, 6.0 % and 5.5 % of the trials were removed, respectively.

A short summary of the analyses will be given in the next paragraph, followed by a detailed method description in the remaining section.

Faster responses to bimodal, compared to unimodal, stimuli indicate an RSE. Therefore, we first compared bimodal to unimodal RTs by calculating the difference between the bimodal and the faster of the two unimodal RTs (visual or olfactory). Because this global RSE is relatively insensitive to effects that are not present across the whole response time range, we next estimated cumulative distribution functions (CDFs) from the RTs. Analyses based on these CDFs can take into account the whole RT distribution. We evaluated the CDFs at 10 quantiles. Since an observed response speedup can be caused by statistical facilitation alone, in a next step we calculated theoretical model boundaries based on the unimodal CDFs under the assumption of parallel processing of the visual and olfactory channels (race model), that is the data range that could be explained by statistical facilitation. Any observation exceeding these limits would support the hypothesis of true integrative processing. To examine whether the channels operated in a stochastically independent manner, we additionally compared our data to a very specific race model assuming stochastic independence of the channel processing speeds (UCIP model). A very similar approach was chosen in the analysis of the accuracy data, although it was naturally based on mean accuracies and not single-trial responses, i.e., no equivalent of a CDF could be estimated.

4.4.1. Response Times

The RT distributions were heavily positively skewed; we therefore used the median as measure of central tendency. This measure is not without criticism (cf. Miller, 1988), but alternatives like the commonly applied log-transformations are not universally applicable approaches either (Feng et al., 2014).

Response times to unimodal V and O stimuli and their bimodal VO combination are defined as non-negative random vectors RT_V , RT_O and RT_{VO} . Their respective expected values shall be labeled $E(RT_V)$, $E(RT_O)$ and $E(RT_{VO})$, and their distribution functions as $F(RT_V)$, $F(RT_O)$, and $F(RT_{VO})$. An RSE can be observed if

$$E(RT_{VO}) < \min[E(RT_V), E(RT_O)], \quad (1)$$

i.e., if mean RTs for bimodal VO stimuli are faster than for either unimodal component.

We calculated the difference between the medians of the fastest unimodal and the bimodal RTs, i.e., $\min[E(RT_V), E(RT_O)] - E(RT_{VO})$. Positive values indicate a bimodal speedup, i.e., a facilitation in processing of bimodal as compared to unimodal stimuli. Note that RSEs were calculated separately for each object (banana and lemon) before collapsing and submission to one-sample *t*-tests against zero to identify bimodal facilitation.

To quantify the effect of in perceptually aligning the unimodal constituents of bimodal trials in Experiment 2, we compared the median RTs of the unimodal V and O conditions (collapsed across objects) using paired *t*-tests for each experiment.

Next, we tested whether the RT distributions fit probability summation models. In bimodal trials, only the marginal distribution $F(RT_{VO})$, but not the distributions of the unimodal constituents $F(RT_V)$ and $F(RT_O)$ can be observed.

Probability summation models critically rely on the assumption of *context-invariance* (Colonius, 1990), which states that the processing speed of a channel is identical in unimodal and bimodal stimulations, that is additional work load on one channel does not influence processing speed in another channel, suggesting *unlimited capacity*.

The unlimited capacity, independent, parallel (UCIP) model makes the additional assumption that the processing speeds of individual channels are uncorrelated and hence *stochastically independent* (Raab, 1962; Meijers and Eijkman, 1977). According to a UCIP model, the cumulative distribution function for the bimodal stimulation is:

$$F(RT_{VO})(t) = F(RT_V)(t) + F(RT_O)(t) - F(RT_V)(t) \times F(RT_O)(t). \quad (2)$$

The last term is always equal to or greater than zero, i.e., $F(RT_V)(t) \times F(RT_O)(t) \geq 0$.

Miller (1982) discarded the assumption of stochastic independence and instead assumed a maximally negative dependence between the channel processing speeds (Colonius, 1990). This allowed him derive an upper bound for the maximum achievable performance gain under *any* parallel processing model called *Miller bound* or *race model inequality* (RMI), commonly expressed as:

$$F(RT_{VO})(t) \leq F(RT_V)(t) + F(RT_O)(t) \quad (3)$$

All parallel processing models have to satisfy inequality (3). If the inequality is violated, the assumption of parallel processing must be dropped, i.e., all race models are ruled out immediately, and the results can only be accounted for by what Miller called *coactivation models* (Miller, 1982)¹. Similarly, a lower performance bound was defined by Grice et al. (1984b), implying perfect positive dependence (Colonius, 1990) between the channels' processing speeds:

$$F(RT_{VO})(t) \geq \max[F(RT_V)(t), F(RT_O)(t)] \quad (4)$$

That is, performance in the bimodal conditions should be equal to or faster than in the fastest unimodal condition.

In the case of asynchronous stimulation, i.e., by delaying the presentation of the visual stimulus by the time τ , Equations (1), (2), respectively, become (Miller, 1986):

$$E(RT_{VO(\tau)}) < \min[E(RT_V + \tau), E(RT_O)], \text{ and} \quad (5)$$

$$F(RT_{VO(\tau)})(t) = F(RT_V)(t - \tau) + F(RT_O)(t) - F(RT_V)(t - \tau) \times F(RT_O)(t). \quad (6)$$

Note that the visual RT distribution $F(RT_V)(t - \tau)$ is shifted to the right, which is the correct adjustment for the SOA. The adjusted

¹However, it should be noted that the reverse is not true: Showing that the observations can be described using a parallel processing model does not necessarily exclude coactivation models.

Miller and Grice bounds from Equations (3), (4) can then be expressed as:

$$F(RT_{VO(\tau)})(t) \leq F(RT_V)(t - \tau) + F(RT_O)(t), \text{ and} \quad (7)$$

$$F(RT_{VO(\tau)})(t) \geq \max[F(RT_V)(t - \tau), F(RT_O)(t)]. \quad (8)$$

We estimated empirical cumulative distribution functions (CDFs) of the RTs using a Python implementation of the algorithm suggested by Ulrich et al. (2007). The CDFs predicted by the UCIP model denoted in Equation (6), as well as the theoretical race model boundaries from Equations (7), (8) were calculated based on the unimodal CDFs, resulting in six CDFs per participant (unimodal O and V, bimodal VO, UCIP model, upper and lower bound). All CDFs were then evaluated at ten evenly spaced quantile points (0.05, 0.15, . . . , 0.95), which were subsequently collapsed across both objects. The resulting values were submitted to separate paired *t*-tests for every quantile to test for deviations from the model predictions.

4.4.2. Accuracy

Similar to Equation (1), an RSE in accuracy can be observed if

$$E(ACC_{VO}) > \max[E(ACC_V), E(ACC_O)], \quad (9)$$

i.e., if mean accuracy for bimodal VO stimuli is higher than for the most accurate of the unimodal components.

We calculated the difference between the means of the most accurate unimodal and the bimodal responses, i.e., $E(ACC_{VO}) - \max[E(ACC_V), E(ACC_O)]$. Positive values indicate a bimodal accuracy enhancement. Note that RSEs were calculated separately for each object before collapsing to one-sample *t*-tests against zero.

Following the assumption of the UCIP model, Equation (2) can be applied to accuracy data and becomes (Stevenson et al., 2014):

$$p(ACC_{VO}) = p(ACC_V) + p(ACC_O) - p(RT_V) \times p(ACC_O). \quad (10)$$

Equivalents of the *Miller and Grice bounds* for bimodal accuracy were proposed by Colonius (2015). Formulas (3) and (4), respectively, then become

$$p(ACC_{VO}) \leq p(ACC_V) + p(ACC_O), \text{ and} \quad (11)$$

$$p(ACC_{VO}) \geq \max[p(ACC_V), p(ACC_O)]. \quad (12)$$

The model predictions and boundaries were calculated for each Experiment and object separately. The results were then collapsed across objects. The resulting values were then submitted to paired *t*-tests to test for deviations from the model predictions.

Author Contributions

RH, NB, and KO designed the experiments; RH acquired the data and performed the analyses with input from KO and NB. RH and KO wrote the paper.

Acknowledgments

The authors thank Andrea Katschak for help with data collection and stimulus preparation. Data exploration and aggregation was carried out with IPython (Pérez and Granger, 2007) and pandas (McKinney, 2010). Statistical analyses were conducted with R

(Wickham, 2011; R Core Team, 2014). Plots were generated by ggplot2 (Wickham, 2009) and arranged with Inkscape (<https://inkscape.org/>). The GIMP (<http://www.gimp.org/>) was used for image preparation; the Gaussian blur was generated with SciPy (<http://scipy.org/>). FFT transforms were calculated using FFTW3 (Frigo and Johnson, 2005).

References

- Blechert, J., Meule, A., Busch, N. A., and Ohla, K. (2014). Food-pics: an image database for experimental research on eating and appetite. *Front. Psychol.* 5:617. doi: 10.3389/fpsyg.2014.00617
- Cain, W. S. (1979). To know with the nose: keys to odor identification. *Science* 203, 467–470. doi: 10.1126/science.760202
- Cerf-Ducastel, B., and Murphy, C. (2001). fMRI activation in response to odorants orally delivered in aqueous solutions. *Chem. Sens.* 26, 625–637. doi: 10.1093/chemse/26.6.625
- Chen, K., Zhou, B., Chen, S., He, S., and Zhou, W. (2013). Olfaction spontaneously highlights visual saliency map. *Proc. R. Soc. Lond. B Biol. Sci.* 280:20131729. doi: 10.1098/rspb.2013.1729
- Colonus, H. (1990). Possibly dependent probability summation of reaction time. *J. Math. Psychol.* 34, 253–275. doi: 10.1016/0022-2496(90)90032-5
- Colonus, H. (2015). Behavioral measures of multisensory integration: bounds on bimodal detection probability. *Brain Topogr.* 28, 1–4. doi: 10.1007/s10548-014-0421-3
- Colonus, H., and Vorberg, D. (1994). Distribution inequalities for parallel models with unlimited capacity. *J. Math. Psychol.* 38, 35–58. doi: 10.1006/jmps.1994.1002
- Davis, R. G. (1981). The role of nonolfactory context cues in odor identification. *Percept. Psychophys.* 30, 83–89. doi: 10.3758/BF03206139
- de Araujo, I. E., Rolls, E. T., Kringelbach, M. L., McGlone, F., and Phillips, N. (2003). Taste-olfactory convergence, and the representation of the pleasantness of flavour, in the human brain. *Eur. J. Neurosci.* 18, 2059–2068. doi: 10.1046/j.1460-9568.2003.02915.x
- Diederich, A., and Colonus, H. (2004). Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. *Percept. Psychophys.* 66, 1388–1404. doi: 10.3758/BF03195006
- Feng, C., Wand, H., Lu, N., Chen, T., He, H., Lu, Y., et al. (2014). Log-transformation and its implications for data analysis. *Shanghai Arch. Psychiatry.* 26, 105–109. doi: 10.3969/j.issn.1002-0829.2014.02.009
- Forster, B., Cavina-Pratesi, C., Aglioti, S. M., and Berlucchi, G. (2002). Redundant target effect and intersensory facilitation from visual-tactile interactions in simple reaction time. *Exp. Brain Res.* 143, 480–487. doi: 10.1007/s00221-002-1017-9
- Frigo, M., and Johnson, S. (2005). The design and implementation of FFTW3. *Proc. IEEE* 93, 216–231. doi: 10.1109/JPROC.2004.840301
- Gielen, S. C., Schmidt, R. A., and Van den Heuvel, P. J. (1983). On the nature of intersensory facilitation of reaction time. *Percept. Psychophys.* 34, 161–168. doi: 10.3758/BF03211343
- Girard, S., Collignon, O., and Lepore, F. (2011). Multisensory gain within and across hemispaces in simple and choice reaction time paradigms. *Exp. Brain Res.* 214, 1–8. doi: 10.1007/s00221-010-2515-9
- Gondan, M. (2009). Testing the race model inequality in redundant stimuli with variable onset asynchrony. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 575. doi: 10.1037/a0013620
- Gottfried, J. A., and Dolan, R. J. (2003). The nose smells what the eye sees: crossmodal visual facilitation of human olfactory perception. *Neuron* 39, 375–386. doi: 10.1016/S0896-6273(03)00392-1
- Grice, G. R., and Reed, J. M. (1992). What makes targets redundant? *Percept. Psychophys.* 51, 437–442. doi: 10.3758/BF03211639
- Grice, G. R., and Canham, L. (1990). Redundancy phenomena are affected by response requirements. *Percept. Psychophys.* 48, 209–213. doi: 10.3758/BF03211520
- Grice, G. R., Canham, L., and Boroughs, J. M. (1984a). Combination rule for redundant information in reaction time tasks with divided attention. *Percept. Psychophys.* 35, 451–463. doi: 10.3758/BF03203922
- Grice, G. R., Canham, L., and Gwynne, J. W. (1984b). Absence of a redundant-signals effect in a reaction time task with divided attention. *Percept. Psychophys.* 36, 565–570. doi: 10.3758/BF03207517
- Grill-Spector, K., and Weiner, K. S. (2014). The functional architecture of the ventral temporal cortex and its role in categorization. *Nat. Rev. Neurosci.* 15, 536–548. doi: 10.1038/nrn3747
- Hecht, D., Reiner, M., and Karni, A. (2008). Multisensory enhancement: gains in choice and in simple response times. *Exp. Brain Res.* 189, 133–143. doi: 10.1007/s00221-008-1410-0
- Hershenson, M. (1962). Reaction time as a measure of intersensory facilitation. *J. Exp. Psychol.* 63, 289. doi: 10.1037/h0039516
- Herz, R. S., and von Clef, J. (2001). The influence of verbal labeling on the perception of odors: evidence for olfactory illusions? *Perception* 30, 381–391. doi: 10.1068/p3179
- Hilgard, E. R. (1933). Reinforcement and inhibition of eyelid reflexes. *J. General Psychol.* 8, 85–113. doi: 10.1080/00221309.1933.9713177
- Holmes, N. P. (2007). The law of inverse effectiveness in neurons and behaviour: multisensory integration versus normal variability. *Neuropsychologia* 45, 3340–3345. doi: 10.1016/j.neuropsychologia.2007.05.025
- Laurienti, P. J., Perrault, T. J., Stanford, T. R., Wallace, M. T., and Stein, B. E. (2005). On the use of superadditivity as a metric for characterizing multisensory integration in functional neuroimaging studies. *Exp. Brain Res.* 166, 289–297. doi: 10.1007/s00221-005-2370-2
- Lovelace, C. T., Stein, B. E., and Wallace, M. T. (2003). An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. *Brain Res. Cogn. Brain Res.* 17, 447–453. doi: 10.1016/S0926-6410(03)00160-5
- Lundström, J., Gordon, A. R., Alden, E. C., Boesveldt, S., and Albrecht, J. (2010). Methods for building an inexpensive computer-controlled olfactometer for temporally precise experiments. *Int. J. Psychol.* 78, 179–189. doi: 10.1016/j.ijpsycho.2010.07.007
- McKinney, W. (2010). “Data structures for statistical computing in Python,” in *Proceedings of the 9th Python in Science Conference*, eds S. van der Walt and J. Millman (Austin, TX), 51–56.
- Meijers, L., and Eijkman, E. (1977). Distributions of simple rt with single and double stimuli. *Percept. Psychophys.* 22, 41–48. doi: 10.3758/BF03206078
- Meredith, M. A., and Stein, B. E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *J. Neurophysiol.* 56, 640–662.
- Miller, J. (1982). Divided attention: evidence for coactivation with redundant signals. *Cogn. Psychol.* 14, 247–279. doi: 10.1016/0010-0285(82)90010-X
- Miller, J. (1986). Timecourse of coactivation in bimodal divided attention. *Percept. Psychophys.* 40, 331–343. doi: 10.3758/BF03203025
- Miller, J. (1988). A warning about median reaction time. *J. Exp. Psychol. Hum. Percept. Perform.* 14, 539–543. doi: 10.1037/0096-1523.14.3.539
- O’Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., and Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat. Neurosci.* 4, 95–102. doi: 10.1038/82959
- Olofsson, J. K. (2014). Time to smell: a cascade model of human olfactory perception based on response-time (rt) measurement. *Front. Psychol.* 5:33. doi: 10.3389/fpsyg.2014.00033
- Parker, A., and Gaffan, D. (1998). Memory after frontal/temporal disconnection in monkeys: conditional and non-conditional tasks, unilateral and bilateral frontal lesions. *Neuropsychologia* 36, 259–271. doi: 10.1016/S0028-3932(97)00112-7

- Peirce, J. W. (2009). Generating stimuli for neuroscience using psychopy. *Front. Neuroinformatics* 2:10. doi: 10.3389/neuro.11.010.2008
- Pérez, F., and Granger, B. E. (2007). IPython: a system for interactive scientific computing. *Comput. Sci. Eng.* 9, 21–29. doi: 10.1109/MCSE.2007.53
- Poellinger, A., Thomas, R., Lio, P., Lee, A., Makris, N., Rosen, B. R., et al. (2001). Activation and habituation in olfaction fMRI study. *Neuroimage* 13, 547–560. doi: 10.1006/nimg.2000.0713
- R Core Team. (2014). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Trans. N.Y. Acad. Sci.* 24, 574–590. doi: 10.1111/j.2164-0947.1962.tb01433.x
- Robinson, A. K., Mattingly, J. B., and Reinhard, J. (2013). Odors enhance the saliency of matching images during the attentional blink. *Front. Integr. Neurosci.* 7:77. doi: 10.3389/fnint.2013.00077
- Rolls, E. T., and Baylis, L. L. (1994). Gustatory, olfactory, and visual convergence within the primate orbitofrontal cortex. *J. Neurosci.* 14, 5437–5452.
- Seo, H.-S., Arshamian, A., Schemmer, K., Scheer, I., Sander, T., Ritter, G., et al. (2010). Cross-modal integration between odors and abstract symbols. *Neurosci. Lett.* 478, 175–178. doi: 10.1016/j.neulet.2010.05.011
- Seubert, J., Ohla, K., Yokomukai, Y., Kellermann, T., and Lundström, J. N. (2015). Superadditive opercular activation to food flavor is mediated by enhanced temporal and limbic coupling. *Hum. Brain Mapp.* 36, 1662–1676. doi: 10.1002/hbm.22728
- Shepherd, G. M. (2006). Smell images and the flavour system in the human brain. *Nature* 444, 316–321. doi: 10.1038/nature05405
- Small, D. M., and Green, B. (2012). “A Proposed Model of a Flavor Modality,” in *The Neural Bases of Multisensory Processes*, eds M. M. Murray and M. T. Wallace (Boca Raton, FL: CRC Press), 717–738. doi: 10.1201/9781439812174-46
- Small, D. M., Zald, D. H., Jones-Gotman, M., Zatorre, R. J., Pardo, J. V., Frey, S., et al. (1999). Human cortical gustatory areas: a review of functional neuroimaging data. *Neuroreport* 10, 7–13. doi: 10.1097/00001756-199901180-00002
- Stein, B., and Meredith, M. (1993). *The Merging of the Senses*. Cambridge, MA: MIT Press.
- Stein, B. E., and Meredith, M. A. (1990). Multisensory integration. neural and behavioral solutions for dealing with stimuli from different sensory modalities. *Ann. N.Y. Acad. Sci.* 608, 51–65; discussion: 65–70. doi: 10.1111/j.1749-6632.1990.tb48891.x
- Stevenson, R. A., Ghose, D., Fister, J. K., Sarko, D. K., Altieri, N. A., Nidiffer, A. R., et al. (2014). Identifying and quantifying multisensory integration: a tutorial review. *Brain Topogr.* 27, 707–730. doi: 10.1007/s10548-014-0365-7
- Townsend, J. T., and Nozawa, G. (1995). Spatio-temporal properties of elementary perception: an investigation of parallel, serial, and coactive theories. *J. Math. Psychol.* 39, 321–359. doi: 10.1006/jmps.1995.1033
- Townsend, J. T., and Wenger, M. J. (2004). A theory of interactive parallel processing: new capacity measures and predictions for a response time inequality series. *Psychol. Rev.* 111:1003. doi: 10.1037/0033-295X.111.4.1003
- Ulrich, R., Miller, J., and Schröter, H. (2007). Testing the race model inequality: an algorithm and computer programs. *Behav. Res. Methods* 39, 291–302. doi: 10.3758/BF03193160
- van Wassenhove, V., Grant, K. W., and Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia* 45, 598–607. doi: 10.1016/j.neuropsychologia.2006.01.001
- Veldhuizen, M. G., Shepard, T. G., Wang, M.-F., and Marks, L. E. (2010). Coactivation of gustatory and olfactory signals in flavor perception. *Chem. Senses* 35, 121–133. doi: 10.1093/chemse/bjp089
- Watson, A. B., and Pelli, D. G. (1983). QUEST: A bayesian adaptive psychometric method. *Percept. Psychophys.* 33, 113–120. doi: 10.3758/BF03202828
- Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer.
- Wickham, H. (2011). The split-apply-combine strategy for data analysis. *J. Statist. Softw.* 40, 1–29. doi: 10.18637/jss.v040.i01
- Zellner, D. A., Bartoli, A. M., and Eckard, R. (1991). Influence of color on odor identification and liking ratings. *Am. J. Psychol.* 104, 547–561. doi: 10.2307/1422940
- Zhou, W., Jiang, Y., He, S., and Chen, D. (2010). Olfaction modulates visual perception in binocular rivalry. *Curr. Biol.* 20, 1356–1358. doi: 10.1016/j.cub.2010.05.059
- Zhou, W., Zhang, X., Chen, J., Wang, L., and Chen, D. (2012). Nostril-specific olfactory modulation of visual perception in binocular rivalry. *J. Neurosci.* 32, 17225–17229. doi: 10.1523/JNEUROSCI.2649-12.2012

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Höchenberger, Busch and Ohla. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.