


Emotional Responses to Multisensory Environmental Stimuli: A Conceptual Framework and Literature Review

SAGE Open
January-March 2016: 1–19
© The Author(s) 2016
DOI: 10.1177/2158244016630591
sgo.sagepub.com


Eliane Schreuder¹, Jan van Erp¹, Alexander Toet¹, and Victor L. Kallen¹

Abstract

How we perceive our environment affects the way we feel and behave. The impressions of our ambient environment are influenced by its entire spectrum of physical characteristics (e.g., luminosity, sound, scents, temperature) in a dynamic and interactive way. The ability to manipulate the sensory aspects of an environment such that people feel comfortable or exhibit a desired behavior is gaining interest and social relevance. Although much is known about the sensory effects of individual environmental characteristics, their combined effects are not a priori evident due to a wide range of non-linear interactions in the processing of sensory cues. As a result, it is currently not known how different environmental characteristics should be combined to effectively induce desired emotional and behavioral effects. To gain more insight into this matter, we performed a literature review on the emotional effects of multisensory stimulation. Although we found some interesting mechanisms, the outcome also reveals that empirical evidence is still scarce and haphazard. To stimulate further discussion and research, we propose a conceptual framework that describes how environmental interventions are likely to affect human emotional responses. This framework leads to some critical research questions that suggest opportunities for further investigation.

Keywords

emotion, affective appraisal, environment, multisensory stimuli, cross-modal interactions

Introduction

Sensory input from our environment plays an important role in how we feel and behave (Turley & Milliman, 2000). Although we live in highly diffuse and “vivid” multisensory environments, and despite the growing interest from different application domains, most studies on human *emotional responses* to environmental characteristics still focus on a number of well-defined and restricted sensory aspects of the environment (typically under highly controlled conditions). As a result, we still lack systematic knowledge about successful multisensory interventions that elicit desirable outcomes (Barrett, Barrett, & Davies, 2013; Gerdes, Wieser, & Alpers, 2014; Jain & Bagdare, 2011; Oakes & North, 2008; Spence, Puccinelli, Grewal, & Roggeveen, 2014; Turley & Milliman, 2000). Environmental characteristics such as luminosity of light sources, the nature and level of ambient noise and acoustics, the presence of specific odors, color hues and shades, and materials and atmospheric factors such as temperature and humidity, all generate sensory input, and combined contribute to specific reactions in the observer (Biggers & Pryer, 1982; Franz, 2006). Research from environmental psychology traditionally focused on single characteristics and independent effects of any given sensory modality, such as vision, audition, olfaction, or touch

(Krishna, 2012). However, it is evident from gestalt principles that the sensory input from the environment is not simply perceived as the sum of its individual components, but rather as a whole (Lin, 2004). Experiments conducted in laboratory settings show that there is a broad spectrum of non-linear interactions between all sensory modalities (Bresciani et al., 2005; Demattè, Sanabria, Sugarman, & Spence, 2006; Driver & Noesselt, 2008; Seigneuric, Durand, Jiang, Baudouin, & Schaal, 2010; Shimojo & Shams, 2001; Small, 2004; Thesen, Vibell, Calvert, & Österbauer, 2004). This means that when cues from different sensory modalities are integrated, the result is not a simple accumulation of the effects generated by each modality separately. Main and interaction effects are dynamically intertwined in such a way that effects may be multiplied (sensory cooperation), disambiguated (one cue helps resolve an ambiguity in a second cue), vetoed (a stronger cue is selected over a weaker cue), inhibited, or the stimulation may even lead to an emergent or

¹TNO, Soesterberg, The Netherlands

Corresponding Author:

Alexander Toet, TNO, Kampweg 5, 3769 DE Soesterberg, The Netherlands.
Email: lex.toet@tno.nl



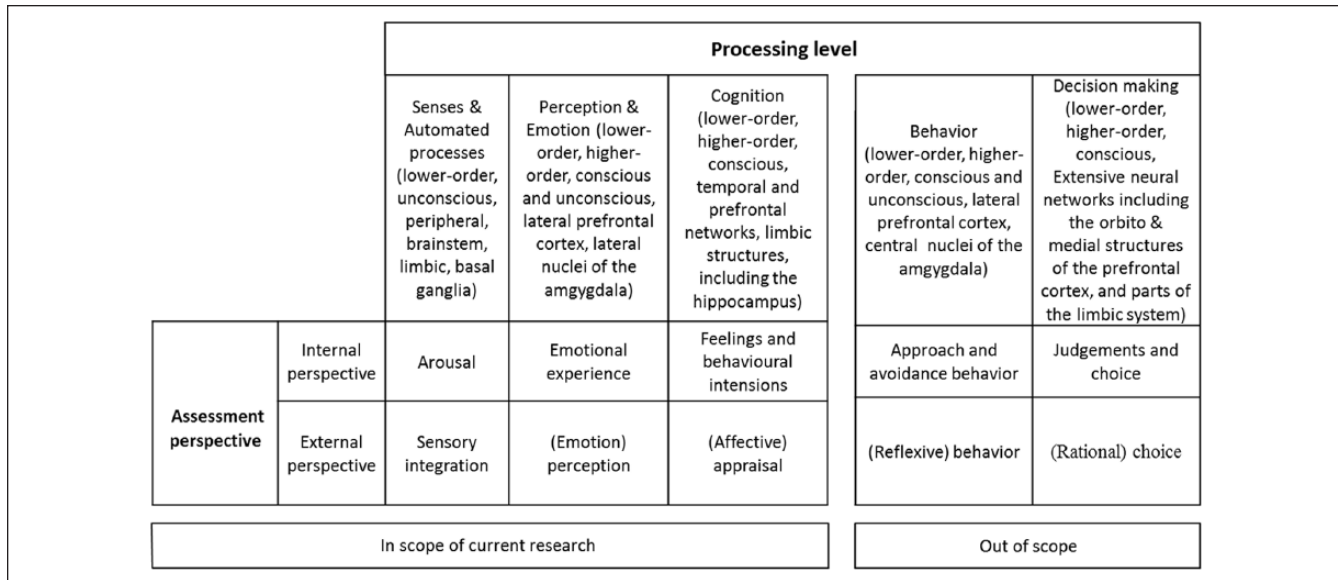


Figure 1. Conceptual multisensory response model.

novel effect (such as the McGurk effect, and illusion that occurs when the auditory component of one sound is paired with the visual component of another sound, leading to the perception of a third sound, McGurk & MacDonald, 1976; or the illusion that a single flash of light is perceived as multiple flashes when it is accompanied by multiple auditory beeps, Shams, Kamitani, & Shimojo, 2002; see also de Gelder & Bertelson, 2003; Gottfried & Dolan, 2004; Helbig & Ernst, 2008; Pourtois, de Gelder, Bol, & Crommelinck, 2005).

Furthermore, the impact of sensory input on, for example, behavior is not only based on sensory cues but also on the social context, personal traits, and mood of the observer. For instance, an excited person perceives odors as more intense (Chen & Dalton, 2005), has a more limited field of view (tunnel vision; Dirkin, 1983), and perceives sounds more selectively (Simoens et al., 2007). People on deserted railway platforms feel safer when light intensities are high and when stimulating music is played, whereas on crowded platforms the same measures increase stress levels (van Hagen, 2011). Also, patients treated in a room with white walls (compared with green walls) disclose more information and have more faith in their practitioner, whereas rooms with white walls may increase patients' stress levels (Dijkstra, 2009). Hence, interventions that induce a desired effect in one environment may have less—or even counterproductive—effects in another environment. Moreover, the same intervention may even have different effects on different populations. This makes it difficult to outline sensory interventions that consistently elicit the desirable emotional or behavioral response over changing or differentiated individual states. Although neurobiological studies have shown that emotional signals delivered via different sensory modalities interact at multiple *processing levels* in the brain, influence each other, and form *holistic percepts*, involving a variety of

brain structures from unisensory cortices to high-level association areas (Klasen, Kreifelts, Chen, Seubert, & Mathiak, 2014), it is still not clear how multisensory input interacts with emotion and behavior.

For this reason, we set out to review the state of the art in research on effects of multisensory stimulation and how multisensory environmental interventions may affect perception and behavior. This study focuses on *emotional* responses, as it is assumed that these (whether consciously perceived or not) are closely linked to behavioral intentions and cognition (Inzlicht, Bartholow, & Hirsh, 2015; Mehrabian & Russell, 1974). Because there is not much literature on the effects of different environmental characteristics on human emotions and behavior in naturalistic settings (Barrett et al., 2013; Jain & Bagdare, 2011; Oakes & North, 2008; Spence et al., 2014; Turley & Milliman, 2000), evidence from laboratory studies are included in this overview as well. To enable a categorization of the found effects, and thereby make it possible to adequately compare, evaluate, and discuss published and future studies on this theme, we propose a conceptual framework in the next section.

Conceptual Framework

In relation to the effects of sensory impact on emotional state, the literature uses a plethora of terms. To order and link the experimental results, we introduce the conceptual framework shown in Figure 1. This framework provides a simplified description of the levels involved in processing (multisensory) stimuli and their link to relevant outcomes, being emotion, cognition, behavior, and decision making. It is based on the environment–human interaction (or Stimulus–Organism–Response) model, introduced by Mehrabian and Russell (1974) and adjusted by Bitner (1992) and Lin (2004).

In this model, the environmental stimuli (S) first evokes an emotional response in individuals (O), which, in turn, potentially elicits either approach or avoidance behavior (R). Two influential models have emerged in the literature, both based on this SOR paradigm. In the first model, emotions (pleasure and arousal) generated by external stimuli have a mediating effect on the appraisal (cognitions) and behavior toward the perceived environment or product. In the second model, which is based on Lazarus cognitive theory of emotions (Lazarus, 1991), emotion has a mediating effect on the relation between appraisal and behavior. Both models have received empirical support in the literature (Fiore & Kim, 2007). We used the first model to describe more closely how multisensory environmental stimuli might be processed and assessed. Whereas previous studies typically differentiate between sensory modalities, our framework is built on two different dimensions: *assessment perspective* and *processing level*.

We distinguish five processing levels from sensing the environmental stimuli to higher order behavioral responses and decision making. Although a hierarchical order exists to some extent, these processing levels also depend on each other and exert bidirectional and unidirectional influences (Franz, 2005; Meier, Robinson, & Clore, 2004). We distinguish two assessment perspectives, related to the object of focus that is assessed and responded to: the *external perspective* in which individuals only assess and respond to information in their environment, and the *internal perspective* in which the internal reaction of the individual to the environmental information is assessed and responded to. We will use this dimension to relate the many different experimental tasks and associated measurement instrument(s) used in the relevant studies of our literature study. For instance, if a person is asked to describe the *experience* or *feeling* while doing a task, an internally focused assessment and response follows, for example, “*I felt excited, stressed.*” If a person is explicitly asked to provide an *affective evaluation* of an object or environment, an externally focused assessment and response follows, for example, “*This object or environment is attractive, boring.*” Both assessment perspectives tap into different processes as we will discuss next.

Lower Order Processes: Senses and Automated Processes

The first processing steps of environmental stimuli are done through our senses and the primary sensory areas in our brain, being automatically and unconsciously, thus without conscious intervention or interpretation (so-called *lower order processes*). The primary structures involved being lower brainstem networks, diverse limbic structures (e.g., the amygdala interacting with the hippocampus), and the basal ganglia. In both assessment perspectives, this processing level results in the sensation of environmental stimuli. For a comprehensive overview of the human sensory anatomy and

automated processes involved, we refer elsewhere (Blake & Sekuler, 2005). In these early processing stages one can, however, already distinguish different processing routes, which are later linked to the assessment perspective (Brosch & Sander, 2013; Pessoa & Adolphs, 2010). One route, that goes through the sensory cortices where feature extraction and *sensory integration* take place, serves to guide the external focus and performs an assessment of environmental stimuli (“external assessment perspective”: Figure 1). At this stage and processing level, the subtle interplay of lower order and top down processes, steering attention and resource allocation, comes in to play (Bishop, 2008; Pessoa, Kastner, & Ungerleider, 2002). This integrative process is supported by a secondary route via the limbic structures (prominently including the amygdala) that affects *arousal* level and influences the internal assessment (“internal assessment perspective”). Efferent networks incorporating the central nuclei of the amygdala and parts of the lateral prefrontal cortex initiate behavioral responses through interaction with afferent trajectories (e.g., sensory pathways from the thalamus) running via lateral nuclei of the amygdala, which are sensitive to valence and mood state. Thus, affecting arousal level is closely associated with prioritizing available processing resources, and setting “the state of mind” and receptiveness (threshold) of the individual for new information (Beck & Clark, 1997). This happens in a dynamic and reciprocal way, with a central role for the amygdala (see Bishop, 2008, as well).

Higher Order Processes: Perception and Emotion

Accumulating neuroimaging research suggests that affective processing involves the interactions of large neural networks in complex, recursive multilevel processes (Brosch & Sander, 2013; Pessoa & Adolphs, 2010). In addition to the automated lower order processes, *higher order processes* (including, for example, previous experiences, information stored in memory) are involved through the hippocampus and temporal cortical structures to integrate and perceive (i.e., make sense of, applying gestalt principles to) the sensory information (O’Callaghan, 2012). The influence of higher order processes depends on factors such as attention and the processing capacities of the individual at that time. This processing level involves conscious as well as unconscious processing. From the external assessment perspective, the integration and interpretation of the sensory information results in a *holistic percept* of an object or environment (e.g., Barrett et al., 2013), whereas it results in an *emotional experience* from the internal assessment perspective. We define an emotional experience or emotion as a short-term state that is directly related to the environmental stimuli. This state (response) is either observed consciously (feeling aroused, pleasant in a specific environment) or unconsciously processed. The (un)conscious emotional experience is then further used as referee for the allocation of processing resources and priorities and affects consecutive processing stadia

(cognition, behavior, and decision) or modulates arousal state (e.g., concentration, attention; Anderson, Siegel, & Barrett, 2011; Zadra & Clore, 2011). Thus, from the external assessment perspective, an observer may for instance perceive a painting or environment with emotional content, and assess it as an emotional scene, but without actually experiencing any emotions. From the internal assessment perspective, an observer may feel arousal and have an *emotional experience* when looking at a scene.

When a high-valenced environmental stimulus is presented (here “valence” refers to the intrinsic attractiveness—positive valence—or averseness—negative valence—of a stimulus) the difference between the two assessment perspectives on this level is as follows. From the external perspective, the interpretation of the emotional qualities of the stimulus (e.g., a fearful object, sad music, a happy human being) results in an *emotion perception*. The internal assessment, however, can result in an *emotional experience* that is evoked in the observer himself or herself by the percept (e.g., “*I feel sad, angry*”). The separation of these perspectives is essential as the perception of emotional qualities is not necessarily accompanied by a consciously perceived or objectively assessable emotional change or (physiological) reaction in the observer (Evans & Schubert, 2008; Gabrielsson, 2002; Kallinen & Ravaja, 2006; Russell & Snodgrass, 1987). Although it was found that for instance music-induced experienced emotions and perceived emotions in response to happy and sad music are highly correlated, it is not clear whether this also holds for emotions induced by stimuli originating from other sensory domains (Konecni, 2008; Scherer, 2004; Zentner, Grandjean, & Scherer, 2008).

Cognition

Once the emotional experience or emotion perception reaches a conscious stage, higher order processes may be involved for cognitive processing. From the external assessment perspective, the primary outcome is an evaluation or *appraisal* of the perceived percept. Depending on the task, this appraisal can be emotional (like or dislike of percept) or functional (evaluation of the characteristics of a percept such as strength, size). We will use the term *affective appraisal* (Russell & Lanius, 1984; Russell & Snodgrass, 1987) to refer to emotional appraisals, to make a clear distinction with the emotional response in the internal assessment perspective. Affective appraisals are the attributed emotional or affective qualities, or cognitions about possible object- or place-elicited holistic percepts (Russell & Snodgrass, 1987).

From the internal assessment perspective, the cognitive processing of emotions may result in conscious *feelings or behavioral intentions*, for example, desire to stay, intention to revisit (also defined as action readiness; Frijda, Kuipers, & Ter Schure, 1989). Whereas emotional experiences are short term and often unconscious, we regard feelings or intentions

as conscious and linked to a specific environment. When feelings or intentions become a long-term conscious experience, possibly triggered by environmental stimuli, but actually more free-floating (i.e., not linked to a specific environment), we regard the response as mood (Frijda, 1993). From a neurobiological perspective this cognitive processing is guided by extensive networks involving orbito- and medial prefrontal structures (external assessment perspective) that intensively interact with the already activated networks involving diverse parts of the limbic system (internal assessment perspective, primarily mediated by the hippocampus and the central amygdaloidal structures; Barbas & Zikopoulos, 2006; Bishop, 2008).

Behavior and Decision Making

Emotion and feelings play a central role in the next two processing levels: behavior and decision making (e.g., Damasio, 1994; Frijda, 1986; Frijda et al., 1989; Lerner, Li, Valdesolo, & Kassam, 2015; Zeelenberg, Nelissen, Breugelmans, & Pieters, 2008). The most widely accepted theory posits that emotion directly causes behavior and that its function is to lead the organism to behave in such a way as to deal with the emotional event (e.g., Cosmides & Tooby, 2000; Frijda, 1986). The competing theory (Baumeister, Vohs, Nathan DeWall, & Liqing, 2007) based on a dual-process model distinguishing between “automatic affect”—simple, fast, and often not conscious—and “conscious emotion”—a more complex phenomenon entailing the awareness of subjective experience—argues that only the former shapes behavior directly, whereas emotion affects behavior indirectly, as a feedback system. According to this perspective, conscious emotion influences cognitive processes, which in turn affect decision making and behavior regulation (Matarazzo & Baldassarre, 2015). The processing levels behavior and decision making, therefore, follow the cognition level in our framework, hence a direct link is assumed with the perception and emotion level (“automatic affect”).

From the external assessment perspective, this direct link to (emotion) perceptions may result in automated highly trained *reflexive behavior* (such as breaking for a red traffic light). Although these types of behavior do involve higher order information (you need to know what a red traffic light means), they do not necessarily involve conscious processing: With routine, and over time, such conditioned responses need less and less externally focused (conscious) attention. This is highly beneficial because it means fewer cognitive resources are needed to “do the job.” If cognitive resources are needed to do the job (the route via cognition), more deliberate (externally motivated) behavior is the response. Next to behavior, in the decision-making process level, *appraisals* may trigger executive functions from the external assessment perspective. These functions manage cognitive processes such as working memory, reasoning, and planning (Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004).

Table 1. Literature Search Terms.

Multisensory stimulation	Emotional response	Type of research
Environment	Emotion	Interaction
Environmental stimuli	Affect	Intervention
Senses	Mood	Experiment
Multisensory	Feeling	
Cross-modal	Attitude	
Multimodal	Anxiety and/or sadness, happiness and/or fear and/or disgust	
Olfaction and/or touch and/or vision and/or audition	and/or surprise and/or stress and/or pleasure, and or/arousal	

As an effect, the appraisal and external criteria may be evaluated and a (*rational*) choice may be the result.

From the internal perspective, emotions may elicit rapid and automated behaviors that can only be acted or reacted (when the initial response appears inaccurate) upon and are hard to prevent. This is, for example, reflected in (un)conscious *approach or avoidance behaviors* (direct link). The route via cognition results in a more deliberate approach or avoidance behaviors influenced by anticipated emotion (DeWall, Baumeister, Chester, & Bushman, 2015). Emotions and feelings also constitute potent, pervasive, predictable, sometimes harmful, and sometimes beneficial drivers of decision making. The underlying mechanisms in which important regularities appear, are described by Lerner et al. (2015). *Judgments and choice* are considered as response in this processing level from the internal perspective.

We introduce this framework for two purposes. The first is to structure and interpret experimental results reported in the literature. In the literature, many different experimental manipulations and measurement instruments are used and this makes the use of some kind of structure imperative to identify key processes. Consequently, we think a framework is indispensable to link these heterogeneous data and to infer generalizable conclusions. The second is to lay the foundation for a structural and potentially computational and predictive model of the effects of multisensory environmental stimuli on, for instance, emotions or behavior. Important questions in this context include the following: *Is the sequence of processing levels fixed? Can processing levels be skipped? Are there mediating factors between processing levels and how do these work? Is there cross-talk between both assessment perspectives, and if so, at what level?*

We should emphasize that we only investigated the effects of multisensory stimuli on emotional responses up to cognition (Figure 1). The effects on behavior and decision making are out of the scope of this research, as behavior and decision making are strongly influenced by emotional responses (DeWall et al., 2015; Lerner et al., 2015). Therefore, we consider understanding the effect of multisensory stimuli on emotions as an essential first step in providing insight into effective environmental interventions.

Literature Study

The aim of the literature study was to investigate the emotional effects of multisensory stimulation by ambient environmental features (e.g., lighting, color, sound, scent; Bitner, 1992) and how interventions in the environment can manipulate emotional responses. Electronic searches were carried out using the databases ScienceDirect, PubMed, PsycINFO, and Google Scholar. Search terms used were combinations of terms from the categories described in Table 1.

Furthermore, related articles were searched based on cited references in articles found relevant. The taste sense was excluded as it is difficult to manipulate emotions through environmental interventions via this modality. The search was conducted between May 2012 and August 2015. Included in this review are studies that were performed in the period between 1974 and 2015 and that

- i. deployed interventions involving environmental stimuli that concurrently stimulate two or more senses, or multiple cues presented in consecution (priming);
- ii. investigated interaction effects or relative effects of multisensory cues; and
- iii. investigated the effect of sensory cues on at least one emotional response.

We stress that this study is about *multisensory* stimulation and its emotional effects. The study of van Rompay, Tanja-Dijkstra, Verhoeven, and van Es (2012), for example, that only manipulated visual (*unisensory*) stimuli (i.e., color and layout) was for this reason excluded. Furthermore, an object itself may not be multimodal, but the appraisal of the object in its environment can be multimodal, for instance, when ambient scent and a product visual is manipulated. In that case, the study fits the inclusion criteria.

Arousal, experienced emotions, feelings, mood, and affective appraisals were considered emotional responses. The appraisal of qualitative characteristics of products or cues such as functionality, sharpness, or loudness was not considered as an emotional outcome. Furthermore, concerning the perception and emotion processing level of the framework, this study focused on experienced emotion (internal

perspective) rather than emotion perception (external perspective). As a result, as Table 1 shows, “emotion perception” was not a specific search term. However, emotion perception articles that were found while searching for studies on emotional responses were included to provide insights into this processing level and how it interacts with other levels. Moderators such as personal traits, social context, and emotional state were considered in the context of found evidence, but were not subject of analysis on their own. The search query resulted in an unknown number of hits (not documented) of which 166 met our inclusion criteria. Of these 166, 83 papers were selected based on the abstract, whereas full text screening finally resulted in 70 relevant papers.

Results From Multisensory Studies

Table 2 presents the results of our literature review structured according to the conceptual framework presented in Figure 1. First, we present papers that approach effects of multisensory stimuli on processing levels from the external assessment perspective. Subsequently, the results from the internal assessment perspective are presented. Within our framework, we work from sensation upward in the processing chain. Papers that used physiological measures to assess arousal are included in the “arousal” category. Papers that only measured arousal with self-reports are included in the emotion and perception level. Papers that cover multiple processing levels are presented in the processing level covering the main output variable, but links are discussed. Papers that include output measures from both assessment perspectives are also presented in the dominant assessment perspective. Although the review focuses on the effect of multisensory stimuli on emotional responses, we also present some additional evidence on effects in other processing levels to be able to link and better interpret the results.

External Assessment Perspective

Multisensory integration and (emotion) perception. There is a growing body of laboratory studies investigating multisensory integration and the effect of multisensory stimulation on human perception. The brain integrates multisensory stimuli from the environment to reduce perceptual ambiguity, improve perceptual performance, judge more precisely, and enhance the detection of stimuli (Helbig & Ernst, 2008; Lalanne & Lorenceau, 2004; Philippi, van Erp, & Werkhoven, 2008). In this process, reciprocal relations exist between our senses. This indicates for instance that vision can influence what we hear, touch, and smell, and vice versa. This means that in a multisensory environment, basically each sensory modality is able to affect the observation in another modality (Bresciani et al., 2005; Seigneuric et al., 2010; Shimojo & Shams, 2001; Thesen et al., 2004). As noted before, effects of sensory cues may be multiplied, disambiguated, vetoed,

inhibited, or the stimulation may even lead to an emergent or novel effect (de Gelder & Bertelson, 2003; Gottfried & Dolan, 2004; Helbig & Ernst, 2008; Pourtois et al., 2005).

Research shows that congruent and incongruent cross-modal conditions elicit different cortical activations (Belardinelli et al., 2004; Calvert & Thesen, 2004; Chen, Yeh, & Spence, 2011; Doehrmann & Naumer, 2008; Driver & Noesselt, 2008; Gottfried & Dolan, 2004; O’Callaghan, 2012; Senkowski, Schneider, Foxe, & Engel, 2008; Thurlings, van Erp, Brouwer, Blankertz, & Werkhoven, 2012). Congruent stimuli (temporal, spatial, or semantic/associative) enhance activation in brain regions mediating stable object representations, whereas incongruent stimuli increase activation in regions involved in cognitive control (Watson et al., 2013). As a result, congruency between stimuli from different modalities facilitates perception, whereas incongruency evokes surprise and stimulates explorative behavior (link to internal assessment; Ludden, Schifferstein, & Hekkert, 2009). Furthermore, it is proposed that depending on the task, the integrated percept is not simply dominated by either one or the other sensory modality. Rather, cues from every modality are integrated or combined such that the most reliable percept is generated to accomplish the task (Helbig & Ernst, 2008; Lalanne & Lorenceau, 2004; Ma & Pouget, 2008). Therefore, the context has a considerable influence on how stimuli are perceived.

The perception of cues with emotional content also received increasing attention. As emotion processing is significant to survival (Cannon, 1932), it is experienced more intensely than non-emotional processing as a result of increased arousal activated via the amygdala (Spreckelmeyer, Kutas, Urbach, Altenmüller, & Münte, 2006). Here, there appears to be a link to the internal assessment perspective. Like non-emotional human perception, emotion perception is also enhanced when emotional information from different modalities is congruent (de Gelder, Morris, & Dolan, 2005; Spreckelmeyer et al., 2006). However, irrespective of the valence congruency between the emotional content from different modalities, the amygdala is activated when the content is sufficiently arousing. Interestingly, the activation of the amygdala is attenuated as soon as one sensory channel carries neutral but meaningful information, next to emotional content carried by another channel (Müller et al., 2011; Müller, Cieslik, Turetsky, & Eickhoff, 2012). This suggests a change in set point for any additional stimulation from other sensory modalities.

Also, emotional information from one modality can automatically and unconsciously influence emotion processing in another, especially when affective information in one modality is ambiguous or undefined (Müller et al., 2012; Müller et al., 2011; Rigoulot & Pell, 2012; Seubert et al., 2010). From that perspective, it is not surprising that the multisensory percept is often influenced in an emotional congruent fashion (Boltz, Ebendorf, & Field, 2009; Ebendorf, 2007; Jeong et al., 2011). For instance, sad (happy) faces are perceived sadder (happier) in combination with music that

Table 2. Overview of Literature per Processing Level and Assessment Perspective.

Assessment perspective	Processing level		
	Senses and automated processes (lower order, unconscious)	Perception and emotion (lower order, higher order, conscious, and unconscious)	Cognition (lower order, higher order, conscious)
Internal perspective	Baumgartner, Esslen, & Jäncke, 2006 Baumgartner, Lutz, Schmidt, & Jäncke, 2006 Ebendorf, 2007 Eldar, Ganor, Admon, Bleich, & Hendler, 2007 Schuurink, Houtkamp, & Toet, 2008 Spreckelmeyer, Kutas, Urbach, Altenmüller, & Münte, 2006	Baumgartner, Esslen, & Jäncke, 2006 Baumgartner, Lutz, et al., 2006 Bolivar, Cohen, & Fentress, 1994 Cohen, 1993 Cottet, Plichon, & Lichtle, 2007 Eldar et al., 2007 Fenko & Looock, 2014 Gabrielsson, 2002 Geringer, Cassidy, & Byo, 1996 Lin, 2010 Liu & Jang, 2009 Marin, Gingras, & Bhattacharya, 2012 Mattila & Wirtz, 2001 Michon & Chebat, 2004 Michon, Chebat, & Turley, 2005 Oakes, 2003 Poon & Grohmann, 2014 Ryu & Jang, 2007 Ryu & Jang, 2008 Schifferstein & Tanudjaja, 2004 Schuurink et al., 2008 Spangenberg, Grohmann, & Sprott, 2005 Tajadura-Jiménez, Larsson, Vålljamäe, Västfjäll, & Kleiner, 2010 Vines, Krumhansl, Wanderley, Dalca, & Levitin, 2011 Vines, Krumhansl, Wanderley, & Levitin, 2006	Cottet et al., 2007 Fiore, Yah, & Yoh, 2000 Liu & Jang, 2009 Ludden, Schifferstein, & Hekkert, 2009 Mattila & Wirtz, 2001 Mitchell, Kahn, & Knasko, 1995 Morrin & Chebat, 2005 Morrison, Gan, Dubelaar, & Oppewal, 2011 Ryu & Jang, 2007 Ryu & Jang, 2008 Spangenberg et al., 2005 Wakefield & Baker, 1998
External perspective	Bresciani et al., 2005 Calvert & Thesen, 2004 Chen, Yeh, & Spence, 2011 de Gelder & Bertelson, 2003 Doehrmann & Naumer, 2008 Geringer, Cassidy, & Byo, 1997 Gottfried & Dolan, 2004 Ma & Pouget, 2008 O'Callaghan, 2012 Philippi, van Erp, & Werkhoven, 2008 Seigneuric, Durand, Jiang, Baudouin, & Schaal, 2010 Senkowski, Schneider, Foxe, & Engel, 2008 Shimojo & Shams, 2001 Thurlings, van Erp, Brouwer, Blankertz, & Werkhoven, 2012	Boltz, Ebendorf, & Field, 2009 Jeong et al., 2011 de Gelder, Morris, & Dolan, 2005 Driver & Noesselt, 2008 Henson & Lillford, 2010 Müller et al., 2011 Müller, Cieslik, Turetsky, & Eickhoff, 2012 Pourtois, de Gelder, Bol, & Crommelinck, 2005 Rigoulot & Pell, 2012 Scherer & Larsen, 2011 Seubert et al., 2010 Spreckelmeyer et al., 2006	Balaji, Raghavan, & Jha, 2011 Bosmans, 2006 Carles, Barrio, & de Lucio, 1999 Eroglu, Machleit, & Chebat, 2005 Konecni, 2008 Kuwano, Namba, Komatsu, Kato, & Hayashi, 2001 Michon & Chebat, 2004 Michon et al., 2005 Morinaga, Aono, Kuwano, & Kato, 2003 Russell, 2002 Schuurink et al., 2008 Spangenberg, Sprott, Grohmann, & Tracy, 2006

evokes a sad (happy) emotion. Remarkably, even unconsciously recognized facial expressions (presented to blind field of participants) seem to modulate fear recognition in the

voice (de Gelder, Pourtois, & Weiskrantz, 2002). The effect was not found for emotional pictures, suggesting a more independent and lower order processing of facial expressions.

In addition, negative items are more likely than positive items to bias a multisensory percept (Scherer & Larsen, 2011; Spreckelmeyer et al., 2006). It is suggested that fearful multisensory stimuli integrate more rapidly and automatically as they are regarded to be of more relevance to immediate survival than happy stimuli (de Gelder et al., 2005; Pourtois et al., 2005). Brain research shows that multisensory integration of positive versus negative emotional cues uses different neuro-anatomical substrates: Convergence areas for happy stimuli pairings are mainly situated anteriorly in the left hemisphere, whereas fear pairings are situated anteriorly in the right hemisphere (Pourtois et al., 2005). Phan, Wager, Taylor, and Liberzon (2002) also found different brain regions involved in processing of different emotions. For instance, fear specifically engages the amygdala, and sadness is associated with activity in the subcallosal cingulate.

To summarize, how multisensory stimuli in the environment are integrated and perceived depends on an individual's context and task. Perception is facilitated when multisensory stimuli are congruent and when emotional content is presented. Incongruent stimuli recruit lower order processes (arousal), possibly because they signal potential conflicts that require cognitive control. Perception is biased by stimulus valence and more easily affected by negative stimuli.

Affective appraisal. Research on the multisensory effects on affective appraisals of the environment has been done in the audio–visual, audio–olfaction, and tactile–olfaction/visual domain. Audio–visual research shows that congruent stimuli increase positive appraisal and that incongruent stimuli negatively influence the appraisal of the environment or product. For instance, congruence between sound and images influences preferences (Carles, Barrio, & de Lucio, 1999). Coherent combinations were rated higher than the mean of the component stimuli. Russell (2002) manipulated the plot of a commercial in a way the message was transferred either explicitly or more implicitly in vision and audio. It was found that the congruent commercial (either explicit or implicit in both modalities) was more persuasive (increased positive attitude toward product). The incongruent commercial was better remembered but induced negative feelings toward the product. They suggested that incongruent presentation feels unpleasant and requires more cognitive effort. In addition, adding visual dynamics to a virtual weather setting involving only sounds marginally increased the positive evaluation of the virtual environment, but did not affect experienced pleasure or arousal measured by self-reports or physiological measures (explicitly no link to internal assessment; Schuurink, Houtkamp, & Toet, 2008). Interestingly, when auditory effects were not corroborated visually, an incongruency effect was found resulting in a negative effect on the appreciation of the environment.

Furthermore, in the audio–visual context, visual cues seem to have more weight in the integrated appraisals than

audio cues, when presented together. For instance, in a setting where participants had to rate the pleasantness of the environment (nature vs. traffic) presented as an image, an audio track, or both, it appeared that a scenery with green plants improved the environment rating even if shown as image only (Kuwano, Namba, Komatsu, Kato, & Hayashi, 2001). Scenes with cars gave negative impressions. Visual masking by green plants seemed effective in reducing negative impression of road traffic noise (Kuwano et al., 2001). Morinaga, Aono, Kuwano, and Kato (2003) also found that perceived pleasantness of a virtual water space is more influenced by visual than auditory information, especially when audio and visual cues are perceived more differently (more ambiguous).

In the audio–olfaction domain, the congruency effect on affective appraisal was also found. Michon and Chebat (2004) studied the interaction between music and scents on the affective appraisal of a shopping mall environment, product quality, and emotion, measured by questionnaires. Mall perceptions improved when the arousing qualities of the stimuli were congruent. This occurred when fast arousing music was played in combination with a positive arousing scent (citrus) as opposed to no scent, or when slow arousing music was combined with the arousing odor (incongruent condition). There was no interaction effect between music and scent on shopper's emotion. However, they did find a main effect of slow music on emotion (suggesting a clear link with internal assessment) but not on shopping mall perception. They suggested that slow music fails to stimulate cognitive processes and as a result, fails to directly affect the appraisal of the mall environment. They also found a moderating effect of emotions on mall perception (link to internal assessment).

Mixed support for the congruency effects was found in the olfaction–visual domain. Studies in the marketing and business domain (e.g., Spangenberg, Sprott, Grohmann, & Tracy, 2006) report that when the perceived gender of a scent is congruent with the perceived gender of product offerings, the store, its merchandise, and actual sales are more positively evaluated. Michon, Chebat, and Turley (2005) however found that store environment appraisals are positively affected when environmental stimuli are mildly incongruent. They investigated effects on the affective appraisal of a shopping environment when scent and mall density were varied. They found that positive scents (lavender, citrus) have an effect on mall perception and marginally on emotion (link to internal assessment) that depends on mall density: no effects in low density, a positive effect in medium density, and a negative effect in high density settings. It was argued by Michon et al. (2005) that a moderately incongruent condition (positive scent vs. moderate mall density) increases arousal (but not to an uncomfortable degree; link to internal assessment) leading to a more favorable evaluation of the environment. According to the authors, this may also explain the negative effect in the high density condition (highly

incongruent and therefore uncomfortable). However, it is not evident why ambient scent would not positively moderate shoppers' perceptions and emotions in the low density (congruent) condition. Finally, Bosmans (2006) found that pleasant non-salient ambient scents enhance product evaluation irrespective of their congruency, whereas salient pleasant scents only enhance product evaluation when they are congruent.

In the tactile-olfaction/visual domain, mixed results were also found for the congruency effect. Krishna, Elder, and Caldara (2010) found that semantic congruence between smell and touch significantly enhances haptic appraisals. Touching a smooth paper (feminine) combined with smelling a feminine scent, led to significantly more positive (in a hedonic sense) haptic appraisals than touching the same paper combined with a masculine smell. The same was true for touching a rough paper (masculine) in combination with the masculine smell. In addition, in the multisensory setting in which participants could feel and see tissues, more positive attitudes toward the product, greater purchase intentions (a clear link to internal assessment), attitude certainty, and importance were reported than in the touch-only or vision-only condition (Balaji, Raghavan, & Jha, 2011). They also found that touch appears dominant over vision in the haptic appraisal of paper tissues. Henson and Lillford (2010) found that dominance of a sense is task dependent: Vision is dominant for "warm" evaluations of textures, whereas touch is dominant for "rough" evaluations. Furthermore, unlike Balaji et al. (2011), no interaction effect was found on the appraisal (simple, rough, warm, like, elegant) of the textures that were seen and/or touched (Henson & Lillford, 2010). The response to multisensory stimuli appeared to simply be a weighted average of the response to individual sensory modalities except for the "natural" evaluations. For natural evaluations, significant interactions between vision and touch were found. Interestingly, although touch provided the clearest cue to distinguish between "natural" and "unnatural" evaluations of the textures, a clear tactile cue did not veto an ambiguous visual cue. Henson and Lillford (2010) argued that appraisals of naturalness were violated because the visual and tactile textures were incongruent (clear tactile cue and ambiguous visual cue).

To summarize, affective appraisals seem positively affected by congruent stimuli from different modalities, and negatively affected by incongruent stimuli. However, in the tactile-vision and vision-olfaction domains, this effect does not always occur, and it may also depend on stimulus salience and level of arousal induced by incongruent stimuli. In audio-visual settings, vision seems dominant. In the vision-tactile domain, the dominant sense seems to depend on the evaluation task.

Internal Assessment Perspective

Arousal and emotional experience. Studies that include measures of arousal and emotion when investigating emotional

responses to multisensory stimuli are generally found in the audio-visual domain. Several studies indicate that congruent multisensory stimuli amplify emotions. Brain studies (Baumgartner, Esslen, & Jäncke, 2006; Baumgartner, Lutz, Schmidt, & Jäncke, 2006; Jeong et al., 2011) all show that pairing of pictures and music conveying the same emotion appears to amplify the experience of the viewer (measured by electroencephalography [EEG] or functional magnetic resonance imaging [fMRI]). Marin, Gingras, and Bhattacharya (2012) more closely reviewed the amplification effect of these congruent pairings, and concluded that these effects do not hold for every emotion category. For example, induced fear in the combined condition did not significantly vary from the picture-only condition. The same is true for the emotion perception equivalent (external assessment perspective): The valence of congruent pairings involving either happy or neutral visual and auditory stimuli was indeed more strongly perceived, but this effect did not hold for sad pairings (Spreckelmeyer et al., 2006). In a laboratory setting in which participants could see and/or listen to a musician performing in an exaggerated, inhibited, or standardized way (using bodily and facial expressions to convey emotions), Vines, Krumhansl, Wanderley, Dalca, and Levitin (2011) found that perceived happiness ratings for the music performance with enhanced bodily expression were significantly higher in the combined condition compared with audio only. Remarkably, the effect was not found for negatively valenced performances. Geringer, Cassidy, and Byo (1996, 1997) also found an *additive effect* of multisensory stimulation. They found that relative to the music-alone condition, some audio-visual formats (in which music was accompanied by videos of cinematic scenes) evoke greater emotional involvement than primarily attributed to a composition's tempo, instrumentation, and dynamics.

Other studies investigated the contribution of an individual modality to emotional experience in a multisensory setting. These studies show that audio and vision can both be dominant depending on the context. For instance, Ellis and Simons (2005) manipulated arousal and valence qualities of films and accompanying music, and measured the emotional response by self-reports and through physiological measures. They found that imagery is more dominant in eliciting emotions than music when simultaneously perceived. Furthermore, an additive relationship was found when music and film were presented together. Positive music elicited higher valence ratings for both positive and negative films. The same relation was found for the effect of highly arousing music on low and high arousing films. This additive relationship received mixed support in physiological data. The interaction effect of music valence and arousal was only found in heart rate and skin conductance, respectively, when film content was positive. They suggested (in accordance with Cohen, 1993, and Bolivar, Cohen, & Fentress, 1994) that music is unable to influence the valence and arousal of highly arousing or negative visual stimuli. Hence, these studies indicate

that, although audio is able to interact with the response to a visual cue, vision is typically more dominant in eliciting emotions than auditory input.

Contrary to these results, Vines, Krumhansl, Wanderley, and Levitin (2006) and Vines et al. (2011) found that audio is the dominant sense in eliciting emotions. Visual information could both enhance and dampen the emotional response evoked by listening to music depending on how coherent the emotion was experienced by both modalities (Vines et al., 2006). However, the authors concluded that a sensory modality's contribution to an experience is task dependent, because vision could be more dominant in other tasks (e.g., a continuous judgment of "amount of movement"). This is supported by studies of Baumgartner, Lutz, et al. (2006) and Baumgartner, Esslen, and Jäncke (2006) in which emotional pictures were paired with emotional music. They showed that perceived accuracy of emotional judgment was stronger in the picture-only compared with the music-only condition, but participants reported an increase in emotional involvement in the combined and music-only relative to the picture-only condition. The combined condition showed the greatest activation in a distributed neuronal network for emotion and arousal processing (measured by EEG, and psychophysiological and psychometrical measures). They suggested that emotional pictures evoke a more cognitive mode of emotion perception, whereas congruent presentations of emotional visual and musical stimuli rather automatically evoke strong emotional experiences. It also suggests, however, a stronger contribution of musical stimuli relative to visual stimuli to emotional involvement.

Marin et al. (2012) investigated whether the valence and arousal of music primes (auditory primes consisting of musical excerpts) presented prior to a visual target, instead of concurrently, could influence the emotional response (self-reported ratings) to emotional visual targets. They found that only arousal induced by music primes modulates arousal in response to visual targets, but no such transfer is observed for pleasantness. It was suggested that the influence of pleasant music on visually induced pleasantness is larger in simultaneously presented stimuli than in consecutive presentation. The effect of arousal, however, appears relatively robust for both cross-modal presentation methods.

Eldar, Ganor, Admon, Bleich, and Hendler (2007) investigated the role of content or meaning on audio–visual interaction. They investigated the effect of adding emotional music poor in concrete content (i.e., containing no meaningful information about the real world) to an emotionally neutral film rich in concrete content. They found that the emotional response (observed through fMRI) was stronger in response to the combination of negative music and neutral film clips compared with the same clips presented separately, despite their incongruity. Interestingly, when the emotional music was presented without a film, no such emotional activation was found. These findings strongly suggest that the brain exerts a stronger response to emotional stimuli when these are associated with concrete content.

Tajadura-Jiménez, Larsson, Våljamäe, Västfjäll, and Kleiner (2010) found an emergent emotion as a result of an interaction between audio and vision. In a virtual big room, emotionally neutral sounds were more arousing and more unpleasant than in a virtual small room, and participants had a stronger feeling of an unsafe situation. They also found that natural (as opposed to artificial) sounds are more arousing in larger rooms. Remarkably, no interaction effects were found for negative sounds and room size on arousal.

To summarize, these studies show that multisensory stimulation, especially when positive, can amplify the arousal or emotional response as compared with unimodal stimulation. Both vision and audio can be dominant in eliciting emotions and can influence each other depending on the context. Timing of multisensory stimuli is relevant for cross-modal interactions. Negative cues in a given modality are less likely to be influenced by another modality than positive or neutral cues, whereas the emotional response also depends on the ecological validity of the stimuli.

Feelings and behavioral intentions. Next to studies on arousal and emotions, a number of papers in marketing and consumer behavior research investigated the effect of multisensory information on *behavioral intentions*, either with or without looking at arousal and emotion. Several studies report increased positive effects on behavioral intentions and feelings when multisensory stimuli are congruent. Mattila and Wirtz (2001) looked at the effect of environmental music and scent in a gift shop on consumer emotion, behavior, feelings, and evaluations (external assessment). They found that consumer satisfaction, impulse buying, and approach behavior increase significantly when music and scent have congruent arousal qualities (high vs. low); whereas pleasure scores increase only marginally. Spangenberg, Grohmann, and Sprott (2005) found that the presence of a Christmas scent next to Christmas music led to more favorable store attitudes, stronger intentions to visit, greater pleasure, greater arousal, greater dominance, and more favorable evaluation of the environment (link to external assessment) compared with a no-scent condition. However, when a Christmas scent was added to "other" music (unrelated to Christmas), no effect on pleasure, arousal, or perceived environment (link to external assessment) was found, and it even led to less dominance, less favorable store and merchandise attitudes (external assessment), and weaker visit intentions. In a similar study, Morrison, Gan, Dubelaar, and Oppewal (2011) reported a congruency effect between music and scent: A combination of high volume music and vanilla aroma (congruent stimuli in the sense that they both induced arousal) significantly enhanced pleasure levels of customers in a shopping environment, which in turn positively affected their shopping behavior. However, in a study on the influence of ambient lavender scent and instrumental music (congruent stimuli in the sense that they both scored high on pleasure and low on arousal) on patients' anxiety in a waiting room of a plastic

surgeon, Fenko and Look (2014) found that music and scent separately each reduced patients' anxiety whereas their combined application had no effect. This suggests that the effects of stimulus congruency are context dependent.

Other papers report that congruence perceived between stimuli and the image of a product, store, or display affect consumer experience and decision making. Cottet, Plichon, and Lichtle (2007) found that music, scents, and colors influence feelings when the cues were congruent with the image of the outlet. Fiore, Yah, and Yoh (2000) reported that more positive effects on approach responses and pleasurable experiences were found when a product display was appropriately fragranced (congruent setting) compared with an inappropriately but pleasantly fragranced product display (incongruent setting), the product alone, or the product in the display without an environmental fragrance. Others (Fiore et al., 2000; Mitchell, Kahn, & Knasko, 1995) found that congruence between ambient scents (chocolate/candy store like or flower shop like) and target product class (chocolates or flowers) improved consumer decision making. They suggested that congruency may increase cognitive flexibility as opposed to incongruency of ambient scents and product class.

In addition, congruence between the emotional state of the observer and the environment affects the impact of the environment on emotions. Lin (2010) found that satisfaction in a bar was increased when color and music settings (either tranquil or dynamic) were congruent with the arousal state of the customer. Morrin and Chebat (2005) varied the presence of ambient scent and music in a shopping mall and found that atmospheric cues were more effective at enhancing consumer response when they were congruent with an individual's affectively or cognitively oriented shopping style.

Studies on the interaction of ambient cues and social density (i.e., the number of individuals in a limited space during a given time period) on the response of people in closed spaces show mixed effects of congruency. Oakes (2003) investigated the effects of congruency between music tempo and social density on feelings of stress in an undergraduate registration queue context. He reported that congruous (low arousal) conditions (slow-tempo music and low social density) enhanced feelings of relaxation in a waiting environment. Poon and Grohmann (2014) investigated the impact of crowd density and ambient scent on people's perception of spatial density (i.e., the amount of objects in a limited space) and anxiety. In conditions of high spatial density (a condition that is known to induce tension; Eroglu & Harrell, 1986), they found a positive effect of stimulus incongruency (an ambient scent associated with spaciousness decreased anxiety levels compared with an ambient scent associated with enclosed spaces); and in conditions of low spatial density, they observed a negative effect of stimulus congruency (an ambient scent associated with spaciousness significantly increased participants' anxiety levels, compared with a scent associated with physical enclosedness). Also, Eroglu, Machleit, and Chebat (2005) found that consumer evaluations of a shopping experience were highest with

a moderate degree of incongruency between social density and music tempo. Like Michon et al. (2005), they argued that the novelty of a moderate incongruency probably induced arousal, which mediated favorable evaluations (external perspective). A possible explanation for these findings may be found in Berlyne's (1960) optimal arousal theory, which suggests that the relation between an individual's level of arousal and affective state can be represented by a bell-shaped (inverted-U) function. Individuals usually prefer medium levels of arousal. Stimuli causing extreme (either too high or too low) levels of arousal result in negative affect. This could also explain the results found by Morrin and Chebat (2005) and Fenko and Look (2014).

Other research focused on the relative contribution of environmental factors to emotional and cognitive responses in a specific setting. In restaurant settings, vision seems especially capable of influencing positive emotions (pleasure) and arousal. The ambiance (combination of audio, haptic, and olfaction cues) is able to influence negative emotions as well as positive emotions. The research also shows the mediating effect of emotion on behavioral intentions. For instance, Ryu and Jang (2007, 2008) and Liu and Jang (2009) measured client evaluation of restaurant settings: arousal, emotion, behavior intentions, and perceived value (external perspective). Ryu and Jang (2008) found that employees have the most significant effect on arousal and that facility aesthetics (painting, plants, color, wall décor) influence both arousal and pleasure. Ambiance (music, aroma, temperature) and layout (machinery, equipment, furniture) have a significant influence on pleasure only. No effect of lighting or dining equipment was found. In addition, the results revealed that pleasure and arousal had significant impact on behavioral intentions, and pleasure appeared to be the more influential emotion of the two. Liu and Jang (2009) found that although ambiance (lighting, music, scent, temperature) has the greatest impact on positive emotion, it also has a significant effect on negative emotion. Interior design (furnishing, paintings, table setting), spatial layout (seat space, easiness to move around, dining privacy), and human elements (clothing, professionalism, adequateness) only influence positive emotions. They found that emotions directly influence perceived value (external perspective) and behavioral intentions (intentions to revisit). Positive emotions show a stronger capability in predicting perceived value of the restaurant than negative emotions. Interestingly, perceived value (external perspective) not only functions as the greatest contributor to behavioral intentions but also mediates the relationship between emotional responses and behavioral intentions.

In retail/shopping environments, vision seems the most important modality to elicit emotional intentions and feelings, whereas the results for sound (music) are mixed and the effect of haptic cues differs significantly across persons and situations (Peck & Childers, 2003) and is only evident for negative effects. For instance, Liaw (2007) found that visual elements (interior design, visuals, color, aesthetics) and

employee characteristics (e.g., appearance, number, friendliness, helpfulness) significantly affect emotions in store environments, whereas music has no emotional effects. Wakefield and Baker (1998) found that architectural design has the highest contribution to feelings of excitement, whereas interior design contributes most to the desire to stay. Music and layout have a positive effect on both outcome measures. Remarkably, there is a negative influence of temperature and light. This indicates that people are only aware of these cues when they are uncomfortable.

In the evaluation of a spa, haptic environmental cues (climate and softness of fabric) have the greatest influence on pleasure scores, although visual elements (e.g., color, layout, design, cleanliness) also have a significant effect (Kang, Boger, Back, & Madera, 2011). The authors also found that audio cues have a significant direct effect on buying intention, without the intervention of emotion (arousal or pleasure). Olfaction cues have an effect neither on emotion nor on buying intention. All sensory factors were highly correlated, reflecting the multisensory nature of perception.

To summarize, behavioral intentions and feelings seem positively affected when multisensory stimuli are congruent, and when stimuli and emotional state of the observer or stimuli and overall image of the environment (store, product) are congruent. However, effects of multisensory stimuli seem also related to the level of arousal elicited and may negatively impact behavioral intentions and feelings when the elicited arousal is either too high or too low. The optimal arousal level is context dependent. Incongruent stimuli are more likely to negatively affect behavioral intentions and affective appraisals (external perspective) than emotions. Emotions seem to mediate higher order behavioral intentions and affective appraisals. Internal responses to an environment are not simply dominated by either one or the other sensory modality but are rather context and activity (shopping, relaxing, dining) dependent.

Discussion

Our literature review of the emotional effects of multisensory stimulation and how interventions in the environment may elicit desired responses shows that evidence on multisensory effects is still scarce and haphazard. Evidence stems from marketing, laboratory, and brain research. Consequently there is considerable variation in the experimental conditions, methodologies, and measures used. This makes it hard to relate findings from different studies in a single perspective. The available studies however, generally seem to differentiate in an externally focused or a more internally focused assessment of environments, objects, or individuals. In an effort to bring these together, we proposed a conceptual framework to describe how multisensory environmental interventions may affect human perception, emotion, cognition, and behavior. Although interesting mechanisms have been identified, and some promising theses can be formulated using the presented framework and its background, there is yet insufficient evidence to validate a type of framework as postulated here.

Consequently, the ability to formulate multisensory assumptions on effective interventions is yet only hypothetical. Relevant lessons learned and current gaps in our knowledge are discussed in the next sections.

Are Effects of Multisensory Stimuli Always Larger Than Those of Unisensory Stimuli?

As argued before, the effects of multisensory cues are not a result of simply adding the effects of unisensory cues (de Gelder & Bertelson, 2003; Gottfried & Dolan, 2004; Helbig & Ernst, 2008; Pourtois et al., 2005). An important question is which factors influence the multisensory effects. The available studies strongly suggest that congruency of multiple sensory stimuli is a very relevant factor to enhance emotional, cognitive, and behavioral effects (e.g., Baumgartner, Lutz, et al., 2006; Baumgartner, Valko, Esslen, & Jäncke, 2006; Belardinelli et al., 2004; Calvert & Thesen, 2004; Carles et al., 1999; Chen et al., 2011; Cottet et al., 2007; Doehrmann & Naumer, 2008; Driver & Noesselt, 2008; Gottfried & Dolan, 2004; Krishna et al., 2010; Mattila & Wirtz, 2001; O'Callaghan, 2012; Senkowski et al., 2008; Spangenberg et al., 2005; Thurlings et al., 2012). From an ecological perspective, multisensory congruency reduces stimulus uncertainty, which may explain why congruent (redundant) multisensory information is more quickly processed whereas incongruent (conflicting) multisensory information takes longer and elicits arousal (Gerdes et al., 2014). In general, we can conclude that congruent multisensory cues strengthen each other's effects (especially positive effects) with respect to both the internal and external assessment perspective, and that this effect can be disturbed by an incongruency, that mainly has a negative impact on higher order processing levels (affective appraisals and behavioral intentions). This incongruency can be subtle such as a small difference in timing, location, arousing qualities (low or high arousing), gender qualities (female, masculine), meaning (song unrelated to Christmas, Christmas scent), or even presentation mode (explicit, implicit; e.g., Krishna et al., 2010; Russell, 2002; Schuurink et al., 2008; Spangenberg et al., 2005). From a behavioral perspective, this implies that multisensory effects are not per se preferred over unisensory effects. Multisensory interventions should be applied carefully as an unexpected perceived incongruency or overstimulation (Fenko & Loock, 2014; Morrin & Chebat, 2005) may result in undesired effects. A side effect of incongruent sensory cues is that their processing may require more cognitive resources, potentially leading to more negative assessments but also to better memory.

Is the Sequence of Processing Levels Fixed, or Can Processing Levels Be Skipped?

Because only a few studies in our review incorporated responses in multiple processing levels, this question can currently not be answered. In the internal assessment perspective evidence, it was found that pleasure and arousal

directly influence behavioral intentions (Ryu & Jang, 2008) in accordance with the proposed processing sequence. However, cues have also been found that directly impact higher order behavioral intentions, without explicitly affecting arousal or emotions (Kang et al., 2011; Spangenberg et al., 2005). This could imply that processing levels in the internal assessment perspective can be skipped.

It should be argued though, in accordance with Ryu and Jang (2007) that many studies in the field of marketing, for instance, pay attention to customer satisfaction or affective appraisals of the product or environment without taking emotions into account. Moreover, the prevailing way of measuring emotional experiences is through self-reports (e.g. on Pleasure, Arousal and Dominance scales). These techniques require that emotional experiences are consciously reflected upon. However, emotional experiences can be very subtle (i.e., unconscious) and are, therefore, not always reported although they actually affect appraisal and behavior (e.g., DeWall et al., 2015; Miers, Blöte, Sumter, Kallen, & Westenberg, 2011). This may be regarded as a result of the limitations in methodology and measurement techniques used in the available studies. Thus, effects at higher processing levels may have been moderated by unconscious emotions, but now we simply do not know. Due to methodological restrictions, this mediating effect is generally not observed or reported and the results are only interpreted as a direct effect. Therefore, we plea for future research on the relation between multisensory stimulation and emotional and behavioral responses that more systematically incorporates and measures responses in different processing levels and assessment perspectives. Thereby, unconscious responses can be taken into account, for instance, through physiological (arousal related) assessment methods (e.g., Ellis & Simons, 2005; Schuurink et al., 2008). This will generate more insight in which processing level interventions are most effective to reach a desired effect.

Can Some Stimuli Reach Higher Processing Levels Easier?

It seems that congruent, ecologically valid and emotional stimuli are more likely to evoke effects on higher processing levels. There is an interesting difference between negative and positive stimuli. Negative stimuli seem to be more automatically and rapidly processed than positive stimuli (de Gelder et al., 2005; Pourtois et al., 2005). In addition, congruent negative audio–visual stimuli do not result in an amplified negative response, as opposed to an amplified positive response to congruent positive stimuli (Marin et al., 2012). Also, Tajadura-Jiménez et al. (2010) showed that neutral sounds impacted emotions differently in a large compared with a small room, but such an effect was not found for negative sounds. This seems to imply that a single negative stimulus and a combination of multiple negative cues both

evoke a similar response. This seems only true, however, if the negative cue is ecologically valid (Eldar et al., 2007). Thus, multisensory effects differ for positive and negative emotions in the sense that *additive effects* are predominantly found for positive emotions and *dominating effects* for negative emotions. This may be related to the ecological significance of the information. The costs involved with a missed threat may be large, certainly compared with a false alarm. Therefore, a single negative signal may already result in a behavioral response of the organism. For positive emotions, this may be exactly the other way around. Here, the cost of responding to a false alarm (i.e., inadvertently interpreting a cue as positive) may be higher than that of a missed signal. For instance, misplaced trust in the intentions of another organism may lead to threatening situations. Therefore, converging positive cues may be required to minimize the risk.

Furthermore, it was suggested that to influence higher order processes such as affective appraisal or behavioral intentions, stimuli should be sufficiently arousing. This is supported by the finding that incongruent stimuli, requiring more resources to process, are more likely to influence higher order processing (behavioral intentions and affective appraisal) levels only (Mattila & Wirtz, 2001; Russell, 2002; Schuurink et al., 2008). Interestingly, once cues are sufficiently arousing to influence higher order processing levels, emotions seem unaffected (Michon & Chebat, 2004; Michon et al., 2005).

Are There Mediating Factors Between Processing Levels, and if So, How Do These Work?

Although we have focused on the lower order effect of multisensory intervention, the human response to an environment is a result of both lower order information (sensory input) and higher order information, with a central role for the limbic structures. This means that the human response is not only a function of stimulus patterns but also affected by personal traits, knowledge, expectations, and the initial emotional state of a person (Kuhbandner et al., 2009). These factors need to be considered to determine the thresholds at which, respectively, internally or externally focused responses are evoked. This process is unique for every individual and in every context (Turley & Milliman, 2000). But, the different processing levels (and how they are activated by certain stimuli) are generally appreciated in some kind of hierarchical perspective. We suggest that the different processing levels in both assessment perspectives are to some extent “fluent” and highly interactive. This hypothesis can be supported by the underlying neurobiological processes. Therefore, we encourage research in laboratorial settings to validate this assumption and to investigate the neurobiological mechanisms that are triggered by multisensory stimulation and the individual factors that influence these mechanisms for each processing level.

Are Sensory Modalities Linked to Assessment Perspectives or Processing Levels?

This question is difficult to answer as the majority of papers focused on audio–visual interventions, which makes it hard to infer conclusions to other sensory domains. According to Baumgartner, Esslen, and Jäncke (2006), vision activates a more cognitive mode (judgment accuracy) and, therefore, seems more effective to influence cognitive processing levels, whereas auditory information (music) seems more effective to influence automated processes (emotional involvement). Also, Schifferstein and Desmet (2007), who assessed the contribution of the different senses in product evaluation (external assessment), stated that vision is mainly used for functional (cognitive) judgments. Blocking vision resulted in the largest loss of functional information, increased task difficulty and task duration, and fostered dependency. When touch was blocked, the perceived loss of information was smaller, and participants reported that familiar products felt less like their own. Blocking audition resulted in communication problems and a feeling of being cut off. Blocking olfaction mainly decreased the intensity of the experience, thereby having a main role in influencing the affective domain.

Furthermore, olfactory cues are able to influence valence of affective appraisals of visual stimuli when presented in sequence (Demattè, Osterbauer, & Spence, 2007; Li, Moallem, Paller, & Gottfried, 2007) unlike audio cues (Marin et al., 2012). Royet et al. (2000) suggest that emotionally weighted olfactory stimuli have a superior potency over visual and auditory stimuli in activating the amygdala (a brain structure dominantly involved in emotional processing). Furthermore, unlike visual, auditory, and tactile stimuli, olfactory stimuli are directly connected to brain structures related to interpretation, without interferences or filtering by the thalamus (Kandel, Schwartz, Jessell, Siegelbaum, & Hudspeth, 2012). Therefore, smell has a great potential to influence the emotional response to other modalities, when presented concurrently or in sequence.

Interestingly, in the psychological assessment perspective, Michon and Chebat (2004) suggested that music plays a more important role in affecting consumers' (conscious) emotional states and odor in affecting cognition. Perhaps, the influence of olfactory cues on emotion is unconscious as shown by Li et al. (2007) who found an effect of scent on the likability rating of faces only for participants lacking conscious awareness of the scent.

Tactile cues seem to determine specific cognitive affective appraisals of objects such as tissues and textures. Psychologically, tactile information in terms of climate influences emotions in spa settings. In other settings, temperature evokes emotions only when it reaches uncomfortable levels. More research is needed to investigate how the different modalities (consciously or unconsciously) tap into the different processing steps and assessment perspectives.

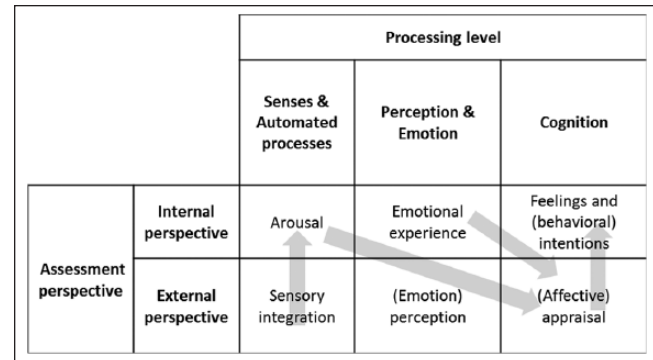


Figure 2. Cross-talks between assessment perspectives.

Is There Cross-Talk Between Both Assessment Perspectives and if So, Where?

Although only a limited number of studies covered multiple output measures from the different assessment perspectives, we did find evidence for cross-talk between them. Figure 2 shows a graphical representation of the observed cross-talks. We found evidence for a link between emotion perception and arousal. When emotional stimuli are processed for emotion perception purposes, arousal levels may automatically increase (Spreckelmeyer et al., 2006). Arousal and emotions can in turn influence affective appraisals. For instance, especially positively experienced emotions or increased arousal can positively influence the affective appraisal of the environment or object (e.g., Liu & Jang, 2009; Michon et al., 2005). However, evidence also shows that affective appraisals can be influenced without intervention of (consciously experienced) arousal or emotions (Michon & Chebat, 2004; Schuurink et al., 2008). Finally, we found that positive affective appraisals (e.g., perceived value of a restaurant) positively influence behavioral intentions, such as intentions to revisit (Liu & Jang, 2009). The same research also found a mediating effect of affective appraisal (perceived value) on the relationship between emotional responses and behavioral intentions.

These observed cross-talks and relations between processing levels and assessment perspectives indicate that human responses are interrelated and bidirectional. We underline that the proposed framework is not aimed at representing the hierarchical order of processing levels, but serves to distinguish commonly used experimental paradigms, to guide research on processes involved in environment–human interaction, and to facilitate the search for effective interventions for desired responses. We believe the conceptual framework is instrumental in this respect, because it forces researchers to unravel these relations, and when well understood, it will provide insight into effective environmental interventions, as these may not simply be directed to the target response level.

Conclusion

Although sensory experiences in the environment can count on broad societal interest, empirical data describing the effect of multisensory input on human emotional, cognitive, and behavioral responses are still scarce and differ in methodology. Combining the results of different multisensory studies however, provided the opportunity to design a gross outline of a multisensory response framework (Figure 1). This framework reflects the way diverse sensory modalities may influence each other's impact on our emotional, cognitive, and behavioral responses. The limited amount of evidence currently available indicates the following:

- Emotional responses to an environment are context dependent and not simply dominated by either one or the other sensory modality.
- Congruency between multiple presented sensory stimuli may enhance emotional, cognitive, and behavioral responses in the positive domain.
- Incongruent multisensory stimuli especially impact and negatively affect higher order responses (affective appraisals and behavioral intentions), but may enhance memory.
- An ecologically valid negative unisensory cue can affect emotions as strongly as a multisensory negative cue.

The body of literature, however, is insufficient to draw conclusions on effective interventions, to validate the proposed framework, or to lay the foundation for a structural and potentially predictive multisensory emotional framework. For these reasons, the proposed framework must be regarded as “conceptual” and instrumental for further discussion and development. To make a meaningful start in this development, we identified gaps in our current empirical knowledge and proposed future research to guide the scientific enterprise with regard to the emotional and behavioral outcomes of multisensory stimulation.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research and/or authorship of this article.

References

- Anderson, E., Siegel, E. H., & Barrett, L. F. (2011). What you feel influences what you see: The role of affective feelings in resolving binocular rivalry. *Journal of Experimental Social Psychology, 47*, 856-860.
- Balaji, M. S., Raghavan, S., & Jha, S. (2011). Role of tactile and visual inputs in product evaluation: A multisensory perspective. *Asia Pacific Journal of Marketing and Logistics, 23*, 513-530.
- Barbas, H., & Zikopoulos, B. (2006). Sequential and parallel circuits for emotional processing in primate orbitofrontal cortex. In D. H. Zald (Ed.), *The orbitofrontal cortex* (Vol. 1, pp. 57-93). Oxford, UK: Oxford University Press.
- Barrett, P., Barrett, L., & Davies, F. (2013). Achieving a step change in the optimal sensory design of buildings for users at all life-stages. *Building and Environment, 67*, 97-104.
- Baumeister, R. F., Vohs, K. D., Nathan DeWall, C., & Liqing, Z. (2007). How emotion shapes behavior: Feedback, anticipation, and reflection, rather than direct causation. *Personality and Social Psychology Review, 11*, 167-203.
- Baumgartner, T., Esslen, M., & Jäncke, L. (2006). From emotion perception to emotion experience: Emotions evoked by pictures and classical music. *International Journal of Psychophysiology, 60*, 34-43.
- Baumgartner, T., Lutz, K., Schmidt, C. F., & Jäncke, L. (2006). The emotional power of music: How music enhances the feeling of affective pictures. *Brain Research, 1075*, 151-164.
- Baumgartner, T., Valko, L., Esslen, M., & Jäncke, L. (2006). Neural correlate of spatial presence in an arousing and noninteractive virtual reality: An EEG and psychophysiology study. *CyberPsychology & Behavior, 9*, 30-45.
- Beck, A. T., & Clark, D. A. (1997). An information processing model of anxiety: Automatic and strategic processes. *Behaviour Research and Therapy, 35*, 49-58.
- Belardinelli, M. O., Sestieri, C., Di Matteo, R., Delogu, F., Del Gratta, C., Ferretti, A., . . . Romani, G. L. (2004). Audio-visual crossmodal interactions in environmental perception: An fMRI investigation. *Cognitive Processing, 5*, 167-174.
- Berlyne, D. E. (1960). *Conflict, arousal, and curiosity*. New York, NY: McGraw-Hill Book.
- Biggers, T., & Pryer, B. (1982). Attitude change: A function of emotion-eliciting qualities of environment. *Personality and Social Psychology Bulletin, 8*, 94-99.
- Bishop, S. J. (2008). Neural mechanisms underlying selective attention to threat. *Annals of the New York Academy of Sciences, 1129*, 141-152.
- Bitner, M. J. (1992). Servicescapes: The impact of physical surroundings on customers and employees. *Journal of Marketing, 56*(2), 57-71.
- Blake, R., & Sekuler, R. (2005). *Perception* (5th ed.). New York, NY: McGraw-Hill Higher Education.
- Bolivar, V. J., Cohen, A. J., & Fentress, J. C. (1994). Semantic and formal congruency in music and motion pictures: Effects on the interpretation of visual action. *Psychomusicology: Music, Mind & Brain, 13*, 28-59.
- Boltz, M. G., Ebendorf, B., & Field, B. (2009). Audiovisual interactions: The impact of visual information on music perception and memory. *Music Perception, 27*, 43-59.
- Bosmans, A. (2006). Scents and sensibility: When do (in)congruent ambient scents influence product evaluations? *Journal of Marketing, 70*(3), 32-43.
- Bresciani, J. P., Ernst, M. O., Drewing, K., Bouyer, G., Maury, V., & Kheddar, A. (2005). Feeling what you hear: Auditory signals can modulate tactile tap perception. *Experimental Brain Research, 162*, 172-180.
- Brosch, T., & Sander, D. (2013). Comment: The appraising brain: Towards a neuro-cognitive model of appraisal processes in emotion. *Emotion Review, 5*, 163-168.
- Calvert, G. A., & Thesen, T. (2004). Multisensory integration: Methodological approaches and emerging principles in the human brain. *Journal of Physiology Paris, 98*, 191-205.

- Cannon, W. B. (1932). *The wisdom of the body*. New York, NY: W.W. Norton.
- Carles, J. L., Barrio, I. L., & de Lucio, J. V. (1999). Sound influence on landscape values. *Landscape and Urban Planning*, *43*, 191-200.
- Chen, D., & Dalton, P. (2005). The effect of emotion and personality on olfactory perception. *Chemical Senses*, *30*, 345-351.
- Chen, Y. C., Yeh, S. L., & Spence, C. (2011). Crossmodal constraints on human perceptual awareness: Auditory semantic modulation of binocular rivalry. *Frontiers in Psychology*, *2*, 212.
- Cohen, A. J. (1993). Associationism and musical soundtrack phenomena. *Contemporary Music Review*, *9*, 163-178.
- Cosmides, L., & Tooby, J. (2000). Evolutionary psychology and the emotions. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (Vol. 2, 2nd ed., pp. 91-115). New York, NY: Guilford Press.
- Cottet, P., Plichon, V., & Lichtle, M.-C. (2007, June). *The compared influence of various perceived environment components: An exploratory study*. The 43th La Londe Conference in Marketing Communications and Consumer Behavior Marseille, IAE Aix Graduate School of Management, Aix-Marseille University, France.
- Damasio, A. R. (1994). *Descartes' error: Emotion, reason and the human brain*. New York, NY: G. P. Putnam.
- de Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends in Cognitive Sciences*, *7*, 460-467.
- de Gelder, B., Morris, J. S., & Dolan, R. J. (2005). Unconscious fear influences emotional awareness of faces and voices. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 18682-18687.
- de Gelder, B., Pourtois, G., & Weiskrantz, L. (2002). Fear recognition in the voice is modulated by unconsciously recognized facial expressions but not by unconsciously recognized affective pictures. *Proceedings of the National Academy of Sciences*, *99*, 4121-4126.
- Demattè, M. L., Osterbauer, R., & Spence, C. (2007). Olfactory cues modulate facial attractiveness. *Chemical Senses*, *32*, 603-610.
- Demattè, M. L., Sanabria, D., Sugarman, R., & Spence, C. (2006). Cross-modal interactions between olfaction and touch. *Chemical Senses*, *31*, 291-300.
- DeWall, C. N., Baumeister, R. F., Chester, D. S., & Bushman, B. J. (2015). How often does currently felt emotion predict social behavior and judgment? A meta-analytic test of two theories. *Emotion Review*. Advance online publication. doi:10.1177/1754073915572690
- Dijkstra, K. (2009). *Understanding healing environments: Effects of physical environmental stimuli on patients' health and well-being* (Doctoral thesis). University of Twente, Enschede, The Netherlands.
- Dirkin, G. R. (1983). Cognitive tunneling: Use of visual information under stress. *Perceptual and Motor Skills*, *56*, 191-198.
- Doehrmann, O., & Naumer, M. J. (2008). Semantics and the multisensory brain: How meaning modulates processes of audio-visual integration. *Brain Research*, *1242*, 136-150.
- Driver, J., & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on "sensory-specific" brain regions, neural responses, and judgments. *Neuron*, *57*, 11-23.
- Ebendorf, B. (2007). *The impact of visual stimuli on music perception* (Unpublished BA thesis). Department of Psychology, Haverford College, PA.
- Eldar, E., Ganor, O., Admon, R., Bleich, A., & Hendler, T. (2007). Feeling the real world: Limbic response to music depends on related content. *Cerebral Cortex*, *17*, 2828-2840.
- Ellis, R. J., & Simons, R. F. (2005). The impact of music on subjective and physiological indices of emotion while viewing films. *Psychomusicology: A Journal of Research in Music Cognition*, *19*, 15-40.
- Eroglu, S. A., & Harrell, G. D. (1986). Retail crowding: Theoretical and strategic implications. *Journal of Retailing*, *62*, 347-363.
- Eroglu, S. A., Machleit, K. A., & Chebat, J. C. (2005). The interaction of retail density and music tempo: Effects on shopper responses. *Psychology & Marketing*, *22*, 577-589.
- Evans, P., & Schubert, E. (2008). Relationships between expressed and felt emotions in music. *Musicae Scientiae*, *12*, 75-99.
- Fenko, A., & Looock, C. (2014). The influence of ambient scent and music on patients' anxiety in a waiting room of a plastic surgeon. *HERD: Health Environments Research and Design Journal*, *7*(3), 38-59.
- Fiore, A. M., & Kim, J. (2007). An integrative framework capturing experiential and utilitarian shopping experience. *International Journal of Retail & Distribution Management*, *35*, 421-442.
- Fiore, A. M., Yah, X., & Yoh, E. (2000). Effects of a product display and environmental fragrancing on approach responses and pleasurable experiences. *Psychology & Marketing*, *17*, 27-54.
- Franz, G. (2005). *An empirical approach to the experience of architectural space* (Doctoral thesis). Max Planck Institute for Biological Cybernetics, Tübingen, Germany and the Bauhaus University, Weimar, Germany.
- Franz, G. (2006). Space, color, and perceived qualities of indoor environments. In M. Tolba, S. Soliman, & A. Abdel-Hadi (Eds.), *Environment, health, and sustainable development—Proceedings of the 19th International Association for People-Environment Studies Conference (IAPS 2006)* (pp. 1-8). Seattle, WA: Hogrefe & Huber.
- Frijda, N. H. (1986). *The emotions*. Cambridge, MA: Cambridge University Press.
- Frijda, N. H. (1993). Moods, emotional episodes and emotions. In M. H. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (pp. 381-403). New York, NY: Guilford Press.
- Frijda, N. H., Kuipers, P., & Ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology*, *57*, 212-228.
- Gabrielsson, A. (2002). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae, Special Issue, 2001-2002*, *5*, 123-147.
- Gerdes, A. B. M., Wieser, M. J., & Alpers, G. W. (2014). Emotional pictures and sounds: A review of multimodal interactions of emotion cues in multiple domains. *Frontiers in Psychology*, *5*, Article 1351.
- Geringer, J. M., Cassidy, J. W., & Byo, J. L. (1996). Effects of music with video on responses of nonmusic majors: An exploratory study. *Journal of Research in Music Education*, *44*, 240-251.
- Geringer, J. M., Cassidy, J. W., & Byo, J. L. (1997). Nonmusic majors' cognitive and affective responses to performance and programmatic music videos. *Journal of Research in Music Education*, *45*, 221-233.

- Gottfried, J. A., & Dolan, R. J. (2004). Response to small: Crossmodal integration—Insights from the chemical senses. *Trends in Neurosciences*, *27*, 123-124.
- Helbig, H. B., & Ernst, M. O. (2008). Haptic perception in interaction with other senses. In M. Grunwald (Ed.), *Human haptic perception: Basics and applications* (pp. 235-249). Basel, Switzerland: Birkhäuser.
- Henson, B., & Lillford, S. (2010). *Affective multimodal integration of visual and tactile textures*. Proceedings of the International Conference on Kansei Engineering and Emotion Research KEER2010, Paris, France.
- Inzlicht, M., Bartholow, B. D., & Hirsh, J. B. (2015). Emotional foundations of cognitive control. *Trends in Cognitive Sciences*, *19*, 126-132.
- Jain, R., & Bagdare, S. (2011). Music and consumption experience: A review. *International Journal of Retail & Distribution Management*, *39*, 289-302.
- Jeong, J.-W., Diwadkar, V. A., Chugani, C. D., Sinsoongsud, P., Muzik, O., Behen, M. E., . . . Chugani, D. C. (2011). Congruence of happy and sad emotion in music and faces modifies cortical audiovisual activation. *NeuroImage*, *54*, 2973-2982.
- Kallinen, K., & Ravaja, N. (2006). Emotion perceived and emotion felt: Same and different. *Musicae Scientiae*, *10*, 191-213.
- Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S. A., & Hudspeth, A. J. (2012). *Principles of neural science* (5th ed.). New York, NY: McGraw-Hill.
- Kang, E., Boger, C. A., Back, K.-J., & Madera, J. (2011, January). *The impact of sensory environments on Spagoers' emotion and behavioral intention*. 16th Graduate Students Research Conference. Retrieved from http://scholarworks.umass.edu/gradconf_hospitality/2011/Presentation/77/
- Klasen, M., Kreifelts, B., Chen, Y. H., Seubert, J., & Mathiak, K. (2014). Neural processing of emotion in multimodal settings. *Frontiers in Human Neuroscience*, *8*, Article 822.
- Konecni, V. J. (2008). Does music induce emotion? A theoretical and methodological analysis. *Psychology of Aesthetics, Creativity, and the Arts*, *2*, 115-129.
- Krishna, A. (2012). An integrative review of sensory marketing: Engaging the senses to affect perception, judgment and behavior. *Journal of Consumer Psychology*, *22*, 332-351.
- Krishna, A., Elder, R. S., & Caldara, C. (2010). Feminine to smell but masculine to touch? Multisensory congruence and its effect on the aesthetic experience. *Journal of Consumer Psychology*, *20*, 410-418.
- Kuhbandner, C., Hanslmayr, S., Maier, M. A., Pekrun, R., Spitzer, B., Pasötter, B., & Bäuml, K.-H. (2009). Effects of mood on the speed of conscious perception: Behavioural and electrophysiological evidence. *Social Cognitive and Affective Neuroscience*, *4*, 286-293.
- Kuwano, S., Namba, S., Komatsu, M., Kato, T., & Hayashi, Y. (2001). Auditory and visual interaction in the aesthetic evaluation of environment. *Empirical Studies of the Arts*, *19*, 191-200.
- Lalanne, C., & Lorenceau, J. (2004). Crossmodal integration for perception and action. *Journal of Physiology Paris*, *98*, 265-279.
- Lazarus, R. S. (1991). *Emotion and adaptation*. New York, NY: Oxford University Press.
- Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annual Review of Psychology*, *66*, 799-823.
- Li, W., Moallem, I., Paller, K. A., & Gottfried, J. A. (2007). Subliminal smells can guide social preferences. *Psychological Science*, *18*, 1044-1049.
- Liaw, G.-F. (2007, October). *The influence of multiple store environment cues on shopping mood and patronage satisfaction*. Proceedings of the 7th Global Conference on Business & Economics. Retrieved from http://www.gcbe.us/7th_GCBE/data/confcd.htm
- Lin, I. Y. (2004). Evaluating a servicescape: The effect of cognition and emotion. *International Journal of Hospitality Management*, *23*, 163-178.
- Lin, I. Y. (2010). The combined effect of color and music on customer satisfaction in hotel bars. *Journal of Hospitality Marketing & Management*, *19*, 22-37.
- Liu, Y., & Jang, S. (2009). The effects of dining atmospherics: An extended Mehrabian-Russell model. *International Journal of Hospitality Management*, *28*, 494-503.
- Ludden, G. D. S., Schifferstein, H. N. J., & Hekkert, P. (2009). Visual-tactile incongruities in products as sources of surprise. *Empirical Studies of the Arts*, *27*, 61-87.
- Ma, W. J., & Pouget, A. (2008). Linking neurons to behavior in multisensory perception: A computational review. *Brain Research*, *1242*, 4-12.
- Marin, M. M., Gingras, B., & Bhattacharya, J. (2012). Crossmodal transfer of arousal, but not pleasantness, from the musical to the visual domain. *Emotion*, *12*, 618-631.
- Matarazzo, O., & Baldassarre, I. (2015). Are emotions reliable predictors of future behavior? The case of guilt and other post-action emotions. In S. Bassis, A. Esposito, & F. C. Morabito (Eds.), *Advances in neural networks: Computational and theoretical issues* (pp. 319-328). Cham, Switzerland: Springer.
- Mattila, A. S., & Wirtz, J. (2001). Congruency of scent and music as a driver of in-store evaluations and behavior. *Journal of Retailing*, *77*, 273-289.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746-748.
- Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. Boston, MA: The MIT Press.
- Meier, B. P., Robinson, M. D., & Clore, G. L. (2004). Why good guys wear white: Automatic inferences about stimulus valence based on brightness. *Psychological Science*, *15*, 82-87.
- Michon, R., & Chebat, J.-C. (2004, June). *Service with a citric note: The interaction effect of background music and ambient scent on the perception of service quality*. Proceedings of the 8th International Research Seminar in Service Management, La Londe-les-Maures, France.
- Michon, R., Chebat, J. C., & Turley, L. W. (2005). Mall atmospherics: The interaction effects of the mall environment on shopping behavior. *Journal of Business Research*, *58*, 576-583.
- Miers, A. C., Blöte, A. W., Sumter, S. R., Kallen, V. L., & Westenberg, P. M. (2011). Subjective and objective arousal correspondence and the role of self-monitoring processes in high and low socially anxious youth. *Journal of Experimental Psychopathology*, *2*, 531-550.
- Mitchell, D. J., Kahn, B. E., & Knasko, S. C. (1995). There's something in the air: Effects of congruent and incongruent ambient odor on consumer decision making. *Journal of Consumer Research*, *22*, 229-238.
- Morinaga, M., Aono, S., Kuwano, S., & Kato, T. (2003). Psychological evaluation of waterside space using audio-visual information. *Empirical Studies of the Arts*, *21*, 185-194.

- Morrin, M., & Chebat, J. C. (2005). Person-place congruency: The interactive effects of shopper style and atmospherics on consumer expenditures. *Journal of Service Research*, 8, 181-191.
- Morrison, M., Gan, S., Dubelaar, C., & Oppewal, H. (2011). In-store music and aroma influences on shopper behavior and satisfaction. *Journal of Business Research*, 64, 558-564.
- Müller, V. I., Cieslik, E. C., Turetsky, B. I., & Eickhoff, S. B. (2012). Crossmodal interactions in audiovisual emotion processing. *Neuroimage*, 60, 553-561.
- Müller, V. I., Habel, U., Derntl, B., Schneider, F., Zilles, K., Turetsky, B. I., & Eickhoff, S. B. (2011). Incongruence effects in crossmodal emotional integration. *Neuroimage*, 54, 2257-2266.
- Oakes, S. (2003). Musical tempo and waiting perceptions. *Psychology & Marketing*, 20, 685-705.
- Oakes, S., & North, A. C. (2008). Reviewing congruity effects in the service environment musicscape. *International Journal of Service Industry Management*, 19, 63-82.
- O'Callaghan, C. (2012). Perception and multimodality. In E. Margolis, R. Samuels, & S. P. Stich (Eds.), *Oxford handbook of philosophy and cognitive science* (pp. 92-117). Oxford, UK: Oxford University Press.
- Peck, J., & Childers, T. L. (2003). To have and to hold: The influence of haptic information on product judgments. *Journal of Marketing*, 67(2), 35-48.
- Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala: From a "low road" to "many roads" of evaluating biological significance. *Nature Reviews Neuroscience*, 11, 773-783.
- Pessoa, L., Kastner, S., & Ungerleider, L. G. (2002). Attentional control of the processing of neutral and emotional stimuli. *Cognitive Brain Research*, 15, 31-45.
- Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage*, 16, 331-348.
- Philippi, T. G., van Erp, J. B. F., & Werkhoven, P. J. (2008). Multisensory temporal numerosity judgment. *Brain Research*, 1242, 116-125.
- Poon, T., & Grohmann, B. (2014). Spatial density and ambient scent: Effects on consumer anxiety. *American Journal of Business*, 29, 76-94.
- Pourtois, G., de Gelder, B., Bol, A., & Crommelinck, M. (2005). Perception of facial expressions and voices and of their combination in the human brain. *Cortex*, 41, 49-59.
- Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, 306, 443-447.
- Rigoulot, S., & Pell, M. D. (2012). Seeing emotion with your ears: Emotional prosody implicitly guides visual attention to faces. *PLoS ONE*, 7, e30740.
- Royet, J. P., Zald, D., Versace, R., Costes, N., Lavenne, F., Koenig, O., & Gervais, R. (2000). Emotional responses to pleasant and unpleasant olfactory, visual, and auditory stimuli: A positron emission tomography study. *The Journal of Neuroscience*, 20, 7752-7759.
- Russell, C. A. (2002). Investigating the effectiveness of product placements in television shows: The role of modality and plot connection congruence on brand memory and attitude. *Journal of Consumer Research*, 29, 306-318.
- Russell, J. A., & Lanius, U. F. (1984). Adaptation level and the affective appraisal of environments. *Journal of Environmental Psychology*, 4, 119-135.
- Russell, J. A., & Snodgrass, J. (1987). Emotion and the environment. In D. Stokols & I. Altman (Eds.), *Vol. 1: Handbook of environmental psychology* (pp. 245-280). New York, NY: John Wiley.
- Ryu, K., & Jang, S. (2008). Influence of restaurants' physical environments on emotion and behavioral intention. *The Service Industries Journal*, 28, 1151-1165.
- Ryu, K., & Jang, S. S. (2007). The effect of environmental perceptions on behavioral intentions through emotions: The case of upscale restaurants. *Journal of Hospitality & Tourism Research*, 31, 56-72.
- Scherer, K. R. (2004). Which emotions can be induced by music? What are the underlying mechanisms? And how can we measure them? *Journal of New Music Research*, 33, 239-251.
- Scherer, L. D., & Larsen, R. J. (2011). Cross-modal evaluative priming: Emotional sounds influence the processing of emotion words. *Emotion*, 11, 203-208.
- Schiffstein, H. N. J., & Desmet, P. M. A. (2007). The effects of sensory impairments on product experience and personal well-being. *Ergonomics*, 50, 2026-2048.
- Schiffstein, H. N. J., & Tanudjaja, I. (2004). Visualising fragrances through colours: The mediating role of emotions. *Perception*, 33, 1249-1266.
- Schuurink, E. L., Houtkamp, J. M., & Toet, A. (2008). Engagement and EMG in serious gaming: Experimenting with sound and dynamics in the Levee Patrolter training game. In P. Markopoulos, B. D. Ruyter, W. A. IJsselstein, & D. Rowland (Eds.), *Vol. Lecture Notes in Computer Science 5294. Fun and Games—Proceedings of the 2nd International Conference on Fun and Games* (pp. 139-149). Berlin, Germany: Springer Verlag.
- Seigneuric, A., Durand, K., Jiang, T., Baudouin, J. Y., & Schaal, B. (2010). The nose tells it to the eyes: Crossmodal associations between olfaction and vision. *Perception*, 39, 1541-1554.
- Senkowski, D., Schneider, T. R., Foxe, J. J., & Engel, A. K. (2008). Crossmodal binding through neural coherence: Implications for multisensory processing. *Trends in Neurosciences*, 31, 401-409.
- Seubert, J., Kellermann, T., Loughead, J., Boers, F., Brensinger, C., Schneider, F., & Habel, U. (2010). Processing of disgusted faces is facilitated by odor primes: A functional MRI study. *NeuroImage*, 53, 746-756.
- Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, 14, 147-152.
- Shimojo, S., & Shams, L. (2001). Sensory modalities are not separate modalities: Plasticity and interactions. *Current Opinion in Neurobiology*, 11, 505-509.
- Simoens, V. L., Istok, E., Hyttinen, S., Hirvonen, A., Näätänen, R., & Tervaniemi, M. (2007). Psychosocial stress attenuates general sound processing and duration change detection. *Psychophysiology*, 44, 30-38.
- Small, D. A. (2004). Crossmodal integration—Insights from the chemical senses. *Trends in Neurosciences*, 27, 120-123.
- Spangenberg, E. R., Grohmann, B., & Sprott, D. E. (2005). It's beginning to smell (and sound) a lot like Christmas: The interactive effects of ambient scent and music in a retail setting. *Journal of Business Research*, 58, 1583-1589.

- Spangenberg, E. R., Sprott, D. E., Grohmann, B., & Tracy, D. L. (2006). Gender-congruent ambient scent influences on approach and avoidance behaviors in a retail store. *Journal of Business Research*, *59*, 1281-1287.
- Spence, C., Puccinelli, N. M., Grewal, D., & Roggeveen, A. L. (2014). Store atmospherics: A multisensory perspective. *Psychology & Marketing*, *31*, 472-488.
- Spreckelmeyer, K. N., Kutas, M., Urbach, T. P., Altenmüller, E., & Münte, T. F. (2006). Combined perception of emotion in pictures and musical sounds. *Brain Research*, *1070*, 160-170.
- Tajadura-Jiménez, A., Larsson, P., Våljamäc, A., Västfjäll, D., & Kleiner, M. (2010). When room size matters: Acoustic influences on emotional responses to sounds. *Emotion*, *10*, 416-422.
- Thesen, T., Vibell, J. F., Calvert, G. A., & Österbauer, R. A. (2004). Neuroimaging of multisensory processing in vision, audition, touch, and olfaction. *Cognitive Processing*, *5*, 84-93.
- Thurlings, M. E., van Erp, J. B. F., Brouwer, A.-M., Blankertz, B., & Werkhoven, P. (2012). Control-display mapping in brain-computer interfaces. *Ergonomics*, *55*, 564-580.
- Turley, L. W., & Milliman, R. E. (2000). Atmospheric effects on shopping behavior: A review of the experimental evidence. *Journal of Business Research*, *49*, 193-211.
- van Hagen, M. (2011). *Waiting experience at train stations*. Delft, The Netherlands: Eburon.
- van Rompay, T. J., Tanja-Dijkstra, K., Verhoeven, J. W., & van Es, A. F. (2012). On store design and consumer motivation spatial control and arousal in the retail context. *Environment and Behavior*, *44*, 800-820.
- Vines, B. W., Krumhansl, C. L., Wanderley, M. M., Dalca, I. M., & Levitin, D. J. (2011). Music to my eyes: Cross-modal interactions in the perception of emotions in musical performance. *Cognition*, *118*, 157-170.
- Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, *101*, 80-113.
- Wakefield, K. L., & Baker, J. (1998). Excitement at the mall: Determinants and effects on shopping response. *Journal of Retailing*, *74*, 515-539.
- Watson, R., Latinus, M., Noguchi, T., Garrod, O. G. B., Crabbe, F., & Belin, P. (2013). Dissociating task difficulty from incongruence in face-voice emotion integration. *Frontiers in Human Neuroscience*, *7*, Article 744.
- Zadra, J. R., & Clore, G. L. (2011). Emotion and perception: The role of affective information. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*, 676-685.
- Zeelenberg, M., Nelissen, R. M., Breugelmans, S. M., & Pieters, R. (2008). On emotion specificity in decision making: Why feeling is for doing. *Judgment and Decision Making*, *3*, 18-27.
- Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, *8*, 494-521.

Author Biographies

Eliane Schreuder is an environmental psychology research scientist in behavior and life sciences at The Netherlands Organization for Applied Scientific Research TNO. Her research examines effects of the built environment on human behavior, emotion and cognition, with applications to healthcare and lighting.

Jan van Erp is principal scientist with The Netherlands Organization for Applied Scientific Research TNO and full professor of tangible user interaction at the University of Twente. His research focuses on multisensory perception, cognition and behavior, applied neuroscience, advanced human-computer interaction, augmented reality, and brain-computer interfaces.

Alexander Toet is a senior scientist in the Department of Perceptual and Cognitive Systems at The Netherlands Organization for Applied Scientific Research TNO. His research interests include human sensory perception and the crossmodal interactions between the different sensory modalities.

Victor L. Kallen is psycho-physiologist in the department of Earth, Life, & Social Sciences at The Netherlands Organization for Applied Scientific Research TNO. He is specialized in the neurobiological factors underlying emotional development and functioning. As senior scientist he is involved in multiple programs studying and implementing interventions to manage and relieve emotional states, e.g. in the mental health, and safety & security domains. Additionally he lectures about these topics at Leiden University and VU University Amsterdam.