Stimulus-Response Compatibility of Auditory Stimulus Features: Timbre, Pitch, and Number

Vom Fachbereich Sozialwissenschaften
der Technischen Universität Kaiserslautern
zur Verleihung des akademischen Grades
Doktor der Philosophie (Dr. phil.)
genehmigte

Dissertation

vorgelegt von

Barbara Estner
aus Marktoffingen

Tag der mündlichen Prüfung:	14.04.2014
Dekan:	Prof. Dr. Shanley Allen
Vorsitzender:	Prof. Dr. Rolf Arnold
Gutachter:	1. Prof. Dr. Thomas Lachmann
	2. Prof. Dr. Cees van Leeuwen

D 386

(2014)

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ANOVA	analysis of variance
dER	differences in mean error rates
dRT	differences in mean reaction times
ER	error rates
fMRI	functional magnetic resonance imaging
Hz	hertz (frequency of a sound signal)
IPS	intraparietal sulcus
ISI	inter-stimulus interval
ITI	inter-trial interval
MNL	mental number line
PDP	parallel distributed processing
rmANOVA	ANOVA with repeated measures
RT	reaction times
SMARC	Spatial Musical Associations of Response Codes
SNARC	Spatial Numerical Associations of Response Codes
SPARC	Spatial Pitch Associations of Response Codes
STARC	Strategic Association of Response Codes
SRC	Stimulus-Response Compatibility

1 Introduction

Perceiving objects and interacting with them within our environment determines a great part of human behavior. In order to do so, we have to classify such objects according to the nature of their properties. Thus, we need to be able to identify the values for both qualitative and quantitative parameters of these object properties.

Quantitative measures or magnitudes, respectively, are often unconsciously assessed in order to solve problems. For doing so, internal representations of these magnitudes are required. In the past it has been suggested that such magnitude representations can be imagined as a spatial continuum on which magnitudes are ordered by size. This is true for magnitudes that are easy to identify (e.g., the physical size of an object), as well as for less commonly assessed magnitudes (e.g., the pitch height of a sound) or even abstract magnitudes (e.g., numbers).

During the last 20 years there has been an increasing amount of research on the nature of these internal representations. For instance, to date there is a wide range of studies that have examined number representations, because despite their abstract nature, numbers are easy to identify and compare.

However, objects in our environment typically contain information on more than one stimulus dimension. For example, when we hear a single sound we can identify categorical or qualitative information (e.g., timbre) as well as ordinal or quantitative information (e.g., pitch height, loudness, duration). Therefore, it is important to consider not only isolated object properties, but also the relations between such object or stimulus features and the mechanisms for simultaneous processing of multiple stimulus features.

The present thesis focuses on the question of whether the processing of stimulus features that are represented along differing spatial axes is independent or not. Particularly, it is investigated whether pitch height as a quantitative measure can be viewed independently of other stimulus features. Interferences of pitch processing with another typical auditory characteristic that only contains categorical information (i.e., timbre) are studied as well as interferences with another quantifiable stimulus feature (i.e., number). Therefore, a joint theoretical background is given which provides with details on the isolated and simultaneous processing of such stimulus features as well as with underlying concepts with a special focus on pitch and number. The thesis then reports two new studies in which the question of independent processing of pitch with simultaneously varied stimulus features is examined separately. A joint review on results is given in the general discussion and an outlook on further research follows in the final summary and conclusion.

2 Theoretical Background

This section provides with the theoretical background for the studies that are presented in this thesis. The central concept is that of *Stimulus-Response Compatibility* (SRC, Fitts & Deininger, 1954; Fitts & Seeger, 1953). The compatibility of stimuli and response options affects performance not only in real life settings, but also in experimental choice-reaction tasks (e.g., Fitts & Seeger, 1953; Kornblum, Hasbroucq, & Osman, 1990; Proctor & Vu, 2006). As this concept allows for the investigation of isolated and simultaneous processing of stimulus features, it forms the ground for theoretical considerations in this thesis.

In the beginning, the concept of SRC is defined. Then, an overview on theoretical accounts to explain SRC effects is provided, followed by a more detailed view on the *dimensional overlap model* (Kornblum et al., 1990) and the *principle of polarity correspondence* (Proctor & Cho, 2006; Proctor & Vu, 2006). The *Spatial Numerical Associations of Response Codes* (SNARC; Dehaene, Bossini, & Giraux, 1993) and the *Spatial Musical* (or *Pitch*) *Associations of Response Codes* (SMARC, Rusconi, Kwan, Giordano, & Umiltà, 2006; SPARC, Lidji, Kolinsky, Lochy, & Morais, 2007) effects, which are examples for isolated SRC effects for numerical and non-numerical magnitudes, are then described and accounted for. The section continues with introducing the question of independent processing of simultaneously perceived stimulus features and underlying concepts. In particular, it is considered whether pitch as an isolated stimulus feature can be viewed independently of other, simultaneously perceived features that can evoke SRC effects and that can be task-relevant or not. The theoretical background concludes with a brief overview and the research interest.

2.1 Defining Stimulus-Response Compatibility (SRC)

When stimulus and response in a choice-reaction task share some common features, this will result in enhanced performance (e.g., Fitts & Deininger, 1954; Fitts & Seeger, 1953; Kornblum et al., 1990; Proctor & Vu, 2006). This phenomenon is named the Stimulus-Response Compatibility (SRC, Fitts & Deininger, 1954; Fitts & Seeger, 1953). For instance, when a stimulus presented on the left side of a screen has to be responded to with a left key press (i.e., in a compatible trial), responses will be faster and more accurate than when the response has to be given with a right key press.

A first notion on SRC effects was already made in the 19th century by Donders (1868/1969), who observed that the performance for left or right hand responses to electrical stimulation of the left or right foot was faster when stimulation and response were located on the same side. Almost a hundred years later, the concept of SRC was established by Fitts and colleagues (Fitts & Deininger, 1954; Fitts & Seeger, 1953).

SRC effects have been shown to persist throughout multiple perceptual domains (e.g., Tagliabue, Zorzi, & Umiltà, 2002; Vu, Proctor, & Urcuioli, 2003) and are only dependent on response location, but not on the responding hand (Wallace, 1971). Furthermore, SRC effects are also mostly independent of task. This means that the stimulus feature that causes the overlap of stimulus and response does not have to be task-relevant (Hommel & Prinz, 1997; Kornblum & Lee, 1995; Lu & Proctor, 1995; Proctor & Reeve, 1990; Simon, 1990). Moreover, SRC affects performance even when the compatibility is merely implied (e.g., a very small object is associated with the left side on a continuum representing physical size; Ren, Nicholls, Ma, Chen, & Dyer, 2011; Shaki, Petrusic, & Leth-Steensen, 2012; Shoben, Čech, Schwanenflugel, & Sailor, 1989).

2.2 Mechanisms of Stimulus-Response Compatibility

In the past, a variety of models has been developed in order to account for aspects of SRC (Proctor & Vu, 2006). In the following, a chronological overview on the development of SRC models is given in which these models are briefly described. The selection of models is based on an overview by Proctor and Cho (2006) and expanded by a further theoretical account for SRC. Afterwards, the dimensional overlap model (Kornblum et al., 1990; Kornblum & Lee, 1995) and the principle of polarity correspondence, as established by Proctor and colleagues (Proctor & Cho, 2006; Proctor & Vu, 2006), are introduced and explained in further detail.

2.2.1 Explanatory Approaches for Stimulus-Response Compatibility: An Overview

Generally, models accounting for SRC effects can be allocated within two groups: The group of *conceptual models*, which are stated qualitatively, and the group of *quantitative and computational models*, which allow for the specification of effect magnitudes in addition (Proctor & Vu, 2006). All the accounts described below assume SRC effects to originate on the response selection stage of task solving, rather than on the representational or action planning stage.

The early SRC model of Deininger and Fitts (1956) is a conceptual model which is based on the assumption that there are always one or more steps of information transformation involved in motor performance. The number of steps and the time required for transformations depend on whether a stimulus-response relation "agrees closely with the basic habits or expectancies of individuals, i.e., with individual and with population stereotypes" (Fitts & Seeger, 1953, p. 208). The required type of transformation or recoding determines performance efficiency (Deininger & Fitts, 1956). Specifically, performance becomes more efficient when the number of operations within the transformation decreases (Proctor & Vu, 2006).

Similar to the conceptual SRC model by Deininger and Fitts (1956), Rosenbloom and colleagues (Rosenbloom, 1986; Rosenbloom & Newell, 1987) developed an algorithmic model of SRC in which it is proposed that SRC effects depend on the number of transformational operations which have to be performed. This model presumes that performance in speeded tasks can be viewed as algorithms or programs. By estimating the number of operations that are required to respond to a specific combination of stimulus and response set within an algorithm, it is even possible to determine a rank ordering of reaction times (RT).

In 1990, Kornblum, Hasbroucq, and Osman introduced the *dimensional overlap model*. This model extends the above mentioned ones by the addition of an automatic activation component. While the earlier models focused on intentional transformation processing routes, Kornblum and colleagues (Kornblum et al., 1990; Kornblum & Lee, 1995) added an automatic response-selection route which produces a response associated with a stimulus feature independent of its task-relevance. This response can then interfere with the response chosen according to task instructions. Due to its importance for the present thesis, further details on this theoretical account will be given in the following section (2.2.2).

Connectionist models are models which emphasize the role of short-term and long-term associations instead of a dual-route for response selection (e.g., Barber & O'Leary, 1997; Stoffer & Umiltà, 1997). Specifically, associations of stimuli with responses can either be introduced by task instructions (i.e., short-term associations) or they can be pre-existing and overlearned (i.e., long-term associations). According to connectionist models, SRC depends on the automatic activation and the directness of the short-term associations of stimuli with responses. Connectionist models, like the *parallel distributed processing* (PDP) model developed by Zhang, Zhang, and Kornblum (1999), which is based on the dimensional overlap model, extend the conceptual models, as they also allow for quantitative predictions (Proctor & Vu, 2006).

Hommel, Müsseler, Aschersleben, and Prinz (2001) also brought forth a conceptual account for SRC. Their *theory of event coding* is based on the idea of common coding. It is assumed that within a task, perception and actions are coded in a joint system (Eimer, Hommel, & Prinz, 1995; Hommel et al., 2001, see Figure 1). Action plans and stimulus representations are both composed of feature codes and these cognitive codes, the so-called event codes, are related to external events (Proctor & Vu, 2006). Thus, instead of a translation from sensory to motor code (lower part, unbroken lines), the perception of a stimulus with certain features directly activates actions that produce the same features (upper part, broken lines). This, in turn, results in enhanced performance in case the primed action is also the one that is correct according to task instructions.

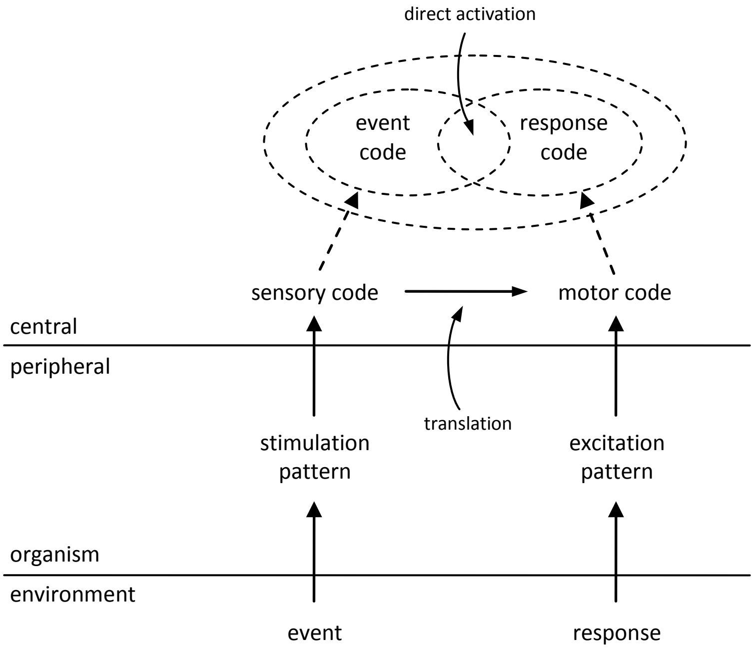

Figure 1: Model of the relationship between perception and action. The lower part depicts the separate coding view involving translation; the upper part (ellipsis) depicts the common coding view. (Eimer et al., 1995)

Another conceptual account is the perspective of *salient features coding* (Proctor & Reeve, 1985; Reeve & Proctor, 1990). This account was initially established for intentional stimulus-response translation and is based on the idea that salient stimulus and response features are involved in translation processes. Consequently, performance is enhanced when salient stimulus and response features systematically correspond (Proctor & Vu, 2006). This kind of correspondence includes three kinds of stimulus features: features, that are intrinsically spatial (e.g., the side of a screen on which a visual stimulus is presented on), features that are represented spatially or ordered in a fixed fashion (e.g., numbers), and features that only allow the asymmetrical forming of categories within this feature (e.g., the parity of a number). Therefore, salient features coding can also account for SRC effects that emerge from a mere structural similarity of stimulus and response set.

The account of salient features coding was further developed by Proctor and colleagues (Proctor & Cho, 2006; Proctor & Vu, 2006) into the principle of *polarity correspondence*. As polarity correspondence is a more general principle that is employed as one of the underlying accounting frameworks in the present thesis, it is described separately and in more detail later (see section 2.2.3).

2.2.2 The Dimensional Overlap Model

In their dimensional overlap model (Figure 2), Kornblum and colleagues (Kornblum et al., 1990; Kornblum & Lee, 1995) were the first ones to establish a non-intentional, automatic route of information processing to account for SRC. The invention of an automatic route allows for the explanation of SRC effects that derive from task-relevant as well as from task-irrelevant stimulus features. The dimensional overlap model assumes two determining factors for SRC: the set-level determinant and the element-level determinant. The set-level determinant, which is termed as dimensional overlap by the authors, forms the representational aspect of the model, while the element-level determinant forms the processing aspect.

Dimensional overlap can be viewed as a consequence of the fact that humans “organize the world along various dimensions and into categories of similar, related, and/or associated objects” (Kornblum et al., 1990, p. 257). Stimulus sets in experimental tasks typically are formed systematically according to specific criteria instead of being chosen randomly, resulting in sets of objects that are homogenous and highly structured. Therefore, object dimensions and attributes are easy to identify which in turn facilitates the transformation of stimuli and responses into classes or categories (Kornblum et al., 1990). If the categories within such sets can be mapped onto each other with a homomorphism (i.e., a transformation that maintains relations and operations within a set, but not necessarily the number of elements), dimensional overlap is present. For instance, one might consider a set of squares (set 1) and a set of circles (set 2) of different sizes, as well as another set of squares (set 3) varying in grayness (from white to black). All three sets contain an order relation (i.e., smaller/larger, brighter/darker grey). Thus, there is dimensional overlap between any two of these three sets. However, dimensional overlap between set 1 and set 2 is greater than for example between set 1 and set 3 due to the greater similarity of the order relations of set 1 and 2 (size) compared to set 1 and 3 (size and grayness).

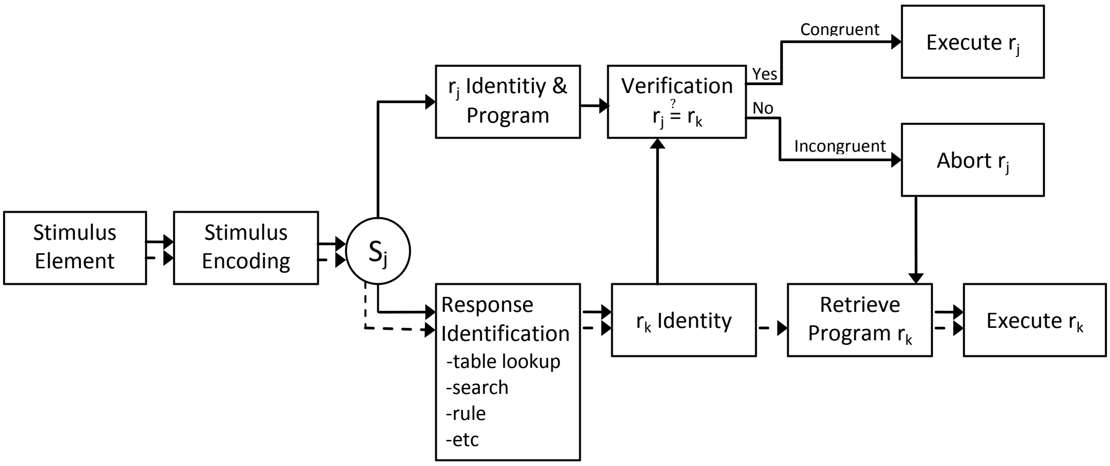

Figure 2: The dimensional overlap model (Kornblum et al., 1990). Continuous lines depict the processing paths when dimensional overlap is present.

When an experimental task is solved, a perceived stimulus is encoded according to its features (set level). The type of coding determines whether dimensional overlap between stimulus and response is present or not. The element-level determinant addresses the processing stage. In any SRC task, the automatic processing route (upper branch) is activated: When a stimulus (S_j) is perceived and dimensional overlap is present (i.e., stimulus and response set share common features in a perceptual, conceptual, or structural manner), the response that is compatible with the stimulus (r_j) is identified and programmed, regardless of whether this stimulus feature is task-relevant or not. In addition, the task-relevant response (r_k) is identified via the intentional route (lower branch), according to task instructions. The time for response identification varies depending on the mapping specified by the instructions. When the mapping follows a rule, this rule can be employed for quick response identification. If, however, no such rule can be applied and stimuli are assigned randomly to responses, the correct response has to be identified by a search through the S-R-table. In both cases, if automatic and identified responses are the same ($r_j = r_k$), the automatically programmed response is carried out; if they are not the same, the automatic response has to be aborted and the motor program for the identified response has to be retrieved and executed. This, in turn, leads to a decrease in performance.

2.2.3 The Principle of Polarity Correspondence

The principle of polarity correspondence has been established by Proctor and colleagues in 2006 (Proctor & Cho, 2006; Proctor & Vu, 2006). It is based on the previously presented account of salient features coding (Proctor & Reeve, 1985; Reeve & Proctor, 1990). The major point of this account is that structural correspondence of stimulus and response sets affects performance in choice reaction tasks. This was for example shown in a study of Proctor and Reeve (1985): The authors observed that when the stimuli O, o, Z, and z had to be mapped to two left- and two right-hand responses that were horizontally aligned, mappings of the type o, O, Z, z produced shorter RT than mappings of the type o, Z, O, z. Proctor and Reeve concluded that the salient feature of the stimulus set was the name of the letter. This categorization by letter name was already sufficient to cause overlap with the salient feature of the response set (i.e., whether a response had to be made with the left or the right hand).

The idea of structural correspondence, as brought forth in the account of salient features coding, also forms the core of the later developed principle of polarity correspondence (Proctor & Cho, 2006). This principle assumes that “for a variety of binary classification tasks, people code the stimulus alternatives and the response alternatives as + polarity and - polarity, and response selection is faster when the polarities correspond than when they do not” (Proctor & Cho, 2006, p. 418). For example, large numbers and right hand responses are both typically coded as + polarity which leads to faster right hand responses for large numbers and faster left hand responses for small numbers within a given set.

It is important to emphasize that polarity correspondence can also be applied to stimulus and response sets that do not contain any perceptual or conceptual overlap. Particularly, any kind of structural similarity of stimulus and response set can lead to consistent polarity coding. This means that associations of stimuli with responses can also occur for stimulus sets that do not contain any spatial or magnitude information. For instance, for number parity it has been shown that participants code odd as - polarity and even as + polarity (Nuerk, Iversen, & Willmes, 2004; Proctor & Cho, 2006).

In sum, polarity correspondence is a comparably simple principle that, similarly to the dimensional overlap account, explains SRC effects for stimulus features when they are intrinsically spatial, when a spatial association (of magnitudes) is only merely implied or even when a feature can only be divided into categories due to a structural asymmetry of the stimulus set concerning that feature. Thus, a variety of effects and interferences in processing stimulus features can be explained within this framework, especially also for the features that are of interest in the present thesis, that is, pitch, timbre, and number.

2.3 Stimulus-Response Compatibility of Isolated Magnitudes: Numbers and Pitch Heights

As implied in earlier sections (sections 2.2.1 and 2.2.3), a stimulus feature does not have to be intrinsically spatial in order to cause SRC. In fact, stimuli that are only considered as activating a mental spatial representation can evoke SRC effects. For example, numbers do not explicitly contain any spatial information but are thought to be represented spatially (for further details, see section 2.3.1.1). Compatibility effects of spatially represented stimuli with responses are said to reflect spatial associations of response codes (Lidji et al., 2007).

Effects of such merely implied spatial associations were first observed for numbers in a study by Dehaene, Bossini, and Giraux in 1993. Since then, a wide range of similar effects have been reported for magnitudes such as physical size (e.g., Ren et al., 2011; Shaki et al., 2012; Shoben et al., 1989), time (e.g., Bonato, Zorzi, & Umiltà, 2012; Ishihara, Keller, Rossetti, & Prinz, 2008a; Ishihara, Keller, Rossetti, & Prinz, 2008b; Santiago, Lupiáñez, Pérez, & Funes, 2007; Vallesi, Binns, & Shallice, 2008), color in synaesthetes (Brugger, Knoch, Mohr, & Gianotti, 2004), luminance (e.g., Cohen Kadosh, Cohen Kadosh, & Henik, 2008b; Pinel, Piazza, Le Bihan, & Dehaene, 2004), pitch (e.g., Cho, Bae, & Proctor, 2012; Lidji et al., 2007; Rusconi et al., 2006), and others (for a review, see Cohen Kadosh, Lammertyn, & Izard, 2008c). Furthermore it has been shown that also the representations of ordinal sequences such as letters of the alphabet or months of the year are spatially organized (e.g., Gevers, Reynvoet, & Fias, 2003).

In the present work, the focus is put to the spatial representation of pitch height and the question, whether SRC effects of pitch height are affected by other stimulus features that can be coded categorically (timbre) or in terms of a magnitude (numbers). Therefore, the following sections introduce the mental representation of numbers and pitch heights and their resulting SRC effects, as for both these magnitudes, SRC is thought to originate from a specific, underlying spatial representation. As numbers were the first magnitude for which such SRC effects were explicitly reported, the focus is first put on this concept, followed by a view on the mental representation of pitch, which is the concept that is centrally investigated in this thesis.

2.3.1 Mental Representation of Numbers

"Number is the within of all things."
(Pythagoras of Samos)

Numbers determine our lives. No matter if written, carved into wood, as knots in a rope or hammered into stone: Numbers have been and continue to be omnipresent in our world and they have been and still are fundamental for the development of modern technology. Therefore, numbers are a cultural invention of great importance and it is impossible to imagine the world without them (Dehaene, 2011).

Considering the importance of numbers in the modern world, it is not surprising that a great deal of research has focused on how this abstract concept is perceived and processed. Numbers were the first magnitude for which an SRC effect due to a merely implied spatial association was reported. The next section therefore gives a short overview on how numbers are assumed to be mentally represented, followed by an introduction to *Spatial Numerical Associations of Response Codes* or, in short, the SNARC effect (Dehaene et al., 1993), and their underlying mechanisms.

2.3.1.1 *Defining Numbers and Their Spatial Representation: The Mental Number Line*

Numbers as a magnitude are special, as they form an abstract magnitude representation (i.e., they represent numerosities). Even though there is a wide range of models accounting for number representation, one of the most accepted ones is the triple-code model established by Dehaene (1992). Dehaene's model (Figure 3) is based on two premises. The first premise states that numbers are mentally represented in three different codes or formats: the *auditory verbal word frame*, the *visual arabic number form* and the *analog magnitude representation*. The second premise states that each numerical procedure is tied to a specific input and output code.

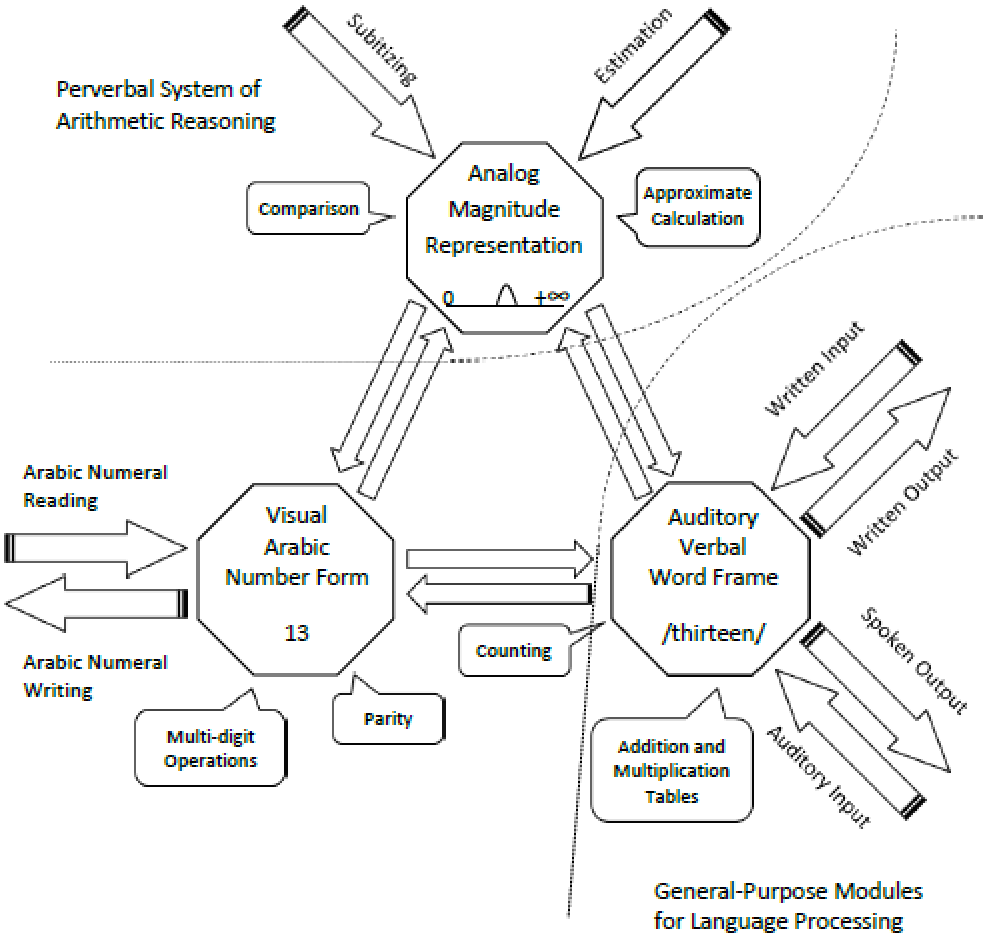

Figure 3: The triple-code model of number processing (Dehaene, 1992)

The auditory verbal word frame is required for processing verbal input. This level of representation processes number words. According to Dehaene (1992), in this format, arithmetical facts (e.g., addition and multiplication tables) and the counting skill are stored. For tasks requiring the retrieval of such facts, the input therefore needs to be transformed to the verbal code before being processed. Visually presented Arabic numerals are processed on the visual arabic number form level. On this level, parity judgments are made and arithmetical operations with multi-digit numbers are carried out. Also for the judgment on whether two numbers are equal or not, they have to first be transformed to the visual arabic format (Dehaene et al., 1993). Within the analog magnitude representation level, finally, the magnitude of a number and its distance relations with other numbers are represented on an oriented, logarithmically compressed number line (Dehaene, 1992). In this number form, magnitude comparisons and approximate calculation, as well as subitizing and numerical estimation take place.

Numbers and numerosities are essential for comparing values of measured entities. While humans are able to perceive numerosities from very early childhood on (e.g., Antell & Keating, 1983; Lourenco & Longo, 2010; Starkey & Cooper, 1980; Starkey, Spelke, & Gelman, 1990), the meaning of exact numbers has to be learned. Thus, in order to perform comparisons of numbers, humans have to learn that numbers represent numerosities and that they are ordered in a fixed fashion. In other words, they have to develop an analog magnitude code for numbers.

The order of numbers is often depicted along a number line on which, at least in Western cultures, numbers are ordered ascending from left to right, likely due to reading habits (e.g., Shaki, Fischer, & Petrusic, 2009). This number line provides with information required to solve for instance retrieval, comparison, or arithmetic tasks (Dehaene, 2011). In order to effectively and efficiently retrieve ordinal and magnitude information, an internal representation of this number line is required which forms along with the learning of numbers and is, therefore, not innate (Núñez, 2011). This so-called *mental number line* (MNL, Restle, 1970) is thought to be subject to developmental change from a logarithmically compressed representation in childhood towards a linear and analogue one (Case & Okamoto, 1996; Laski & Siegler, 2007; Siegler & Booth, 2004). However, there is still debate on whether the MNL observable in adults is exclusively linear or if the compressed one is also present and utilized when needed (e.g., Gallistel & Gelman, 2000). Due to controversial findings on this issue, it may be the case that we use multiple numerical representations (Siegler & Opfer, 2003) which are accessed flexibly, as determined by the specific requirements of a task (Brannon, Wusthoff, Gallistel, & Gibbon, 2001; Lourenco & Longo, 2009; Núñez, Doan, & Nikoulina, 2011; Viarouge, Hubbard, & Dehaene, 2011).

Generally, the MNL is to be viewed as a hypothetical construct for number representations that are spatial in nature. On the MNL, numerical distance is represented as spatial distance and similar number sizes lead to representational overlap (e.g., Göbel, Shaki, & Fischer, 2011). The concept of an MNL is broadly assumed to underlie the associations of numbers with space (e.g., Santens & Gevers, 2008), as reflected by the SNARC effect first reported by Dehaene et al. (1993).

2.3.1.2 Spatial Numerical Associations of Response Codes: The SNARC Effect

The SNARC effect (Dehaene et al., 1993) is probably the best known example for SRC effects that are attributed to an internal spatial representation of magnitudes. When relative numerical magnitude within a set (e.g., small or large) and response location (e.g., left or right) correspond, reactions are made faster and with higher accuracy than when the mapping of numbers to space is not corresponding or compatible, respectively. This association can be observed even when numerical magnitude is not relevant for the task. For instance, in the study of Dehaene et al. (1993), participants were asked to judge the parity of a number (i.e., whether a number is odd or even) and reacted faster when numerical size and response side corresponded.

The SNARC effect is independent of whether response buttons are horizontally, vertically, or diagonally positioned (i.e., *response dimension*) and how the numbers are presented (i.e., *modality*; Nuerk, Wood, & Willmes, 2005). The effect has shown to persist in a considerable number of studies (for a review and meta-analysis, see Wood, Willmes, Nuerk, & Fischer, 2008). However, SNARC seems to be comparably flexible (Pinhas, Pothos, & Tzelgov, 2013) and depending on the relative magnitude of a number within a given set. This means that numbers are coded as small or large in relation to the numerical range employed in a task. For instance, in a task involving numbers ranging from 0-5, the numbers 4 and 5 will be coded as large numbers, while in a number set ranging from 4-9, 4 and 5 will be coded as (relatively) small numbers (Pinhas et al., 2013). Moreover, on the internal scale from 0 to 10, the number 5 holds a special status (Tzelgov, Meier, & Henik, 1992), which is likely because 5 halves the distance between 0 and 10. Due to this, in stimulus sets with numbers ranging from 0 to 10, it is likely best to choose from a numerical range with 5 as an implicit midpoint. This is even more important as the theory of *place coding* (e.g., Verguts & Fias, 2008) assumes that when a number is perceived, not only the representation of that specific number, but also the representation of the neighboring numbers is activated. Thus, when stimulus sets contain numbers very close to each other, relative coding of numbers as large or small may be impaired. Furthermore, certain conditions, as introduced through task instructions (e.g., Bächtold, Baumüller, & Brugger, 1998; Fias, 2001), incompatible spatial mapping or positioning (e.g., Fischer, Mills, & Shaki, 2010; Fischer, Shaki, & Cruise, 2009; Notebaert, Gevers, Verguts, & Fias, 2006; Shaki & Fischer, 2008), and memory requirements or load (e.g., Fischer, Riello, Giordano, & Rusconi, 2013; Herrera, Macizo, & Semenza, 2008; Lindemann, Abolafia, Pratt, & Bekkering, 2008; van Dijck & Fias, 2011; van Dijck, Gevers, & Fias, 2009), can lead to reversion or dilution of SNARC effects.

2.3.1.3 Mechanisms Underlying SNARC

Commonly, the SNARC effect is thought to reflect a horizontal internal mental representation of numbers ascending from left to right (i.e., the MNL). Already Dehaene et al. (1993) attributed their observation of enhanced performance for compatible response mappings to a systematic association of numbers with space (Wood et al., 2008) and concluded that “the representation of number magnitude is automatically accessed during parity judgment of Arabic digits. This representation may be linked to a mental number line (Restle, 1970), because it bears a natural and seemingly irrepressible correspondence with the natural left-right coordinates of external space” (Dehaene et al., 1993, p. 394).

It is often assumed that the SNARC effect emerges due to stimulus processing along a dual-route architecture (e.g., Gevers, Ratinckx, de Baene, & Fias, 2006). This dual-route architecture reflects the major ideas of the dimensional overlap model (Kornblum et al., 1990): When a stimulus is presented, two processing routes are activated in parallel. While the fast and unconditional route automatically programs the response associated with the stimulus according to spatial associations caused by the MNL, the slower, conditional route identifies the (correct) response according to task instructions. When conditional and unconditional responses are identical, a response can be given relatively fast, whereas when this is not the case, the unconditional response needs to be aborted in order to give the correct response (Gevers et al., 2006). SNARC is therefore thought to originate at a response-selection level rather than on earlier levels of stimulus processing (Daar & Pratt, 2008; Keus, Jenks, & Schwarz, 2005; Keus & Schwarz, 2005; Müller & Schwarz, 2007).

In the past there has been a debate on whether the observation of SNARC implies a MNL or not (e.g., Fias & Fischer, 2005; Gevers & Lammertyn, 2005; Santens & Gevers, 2008; Wood et al., 2008). For example, it is argued that SNARC is just the result of a highly overlearned stimulus-response loop (Fitousi, Shaki, & Algom, 2009) or that the culturally formed number representation on which SNARC is based is not necessarily fundamentally spatial (Núñez et al., 2011). Other findings point towards SNARC not only deriving from the response selection stage of processing (e.g., Fischer, Castel, Dodd, & Pratt, 2003; Salillas, El Yagoubi, & Semenza, 2008) or that SNARC is associated with spatial dimensions other than the horizontal one (Gevers & Lammertyn, 2005, Holmes & Lourenco, 2012). These findings, along with the existing debate, give rise to the question whether SNARC can be fully explained with the concepts of MNL in combination with a dual-route processing architecture.

As an alternative account, the principle of polarity correspondence (Proctor & Cho, 2006, see section 2.2.3) further contests the view that the MNL is the sole basis for SNARC effects (e.g., van Dijck, Gevers, Lafosse, & Fias, 2012; Wood et al., 2008). In view of polarity correspondence, the asymmetrical categorization of numbers as (relatively) small or large within a stimulus set is already sufficient to produce polarity codes that then potentially overlap with the polarity codes of responses. Therefore, a MNL would not be required for inducing SNARC effects, because the spatial association is linked to relative magnitude (Ben Nathan, Shaki, Salti, & Algom, 2009).

Concerning the results of their meta-analysis, Wood et al. (2008) state that even though the polarity correspondence framework can account for a wider range of data than the concept of an MNL, there is clear neuropsychological evidence for the use of a mental image (Hoeckner et al., 2008; Vuilleumier, Ortigue, & Brugger, 2004; Zorzi, Priftis, Meneghello, Marenzi, & Umiltà, 2006; Zorzi, Priftis, & Umiltà, 2002). Wood and colleagues reason that "both coordinated [according to the concept of MNL] and categorical [according to polarity correspondence] spatial relations may be important for the SNARC effect, in particular, and for the association between numbers and space, in general." (Wood et al., 2008, p. 511).

Interestingly, the dimensional overlap model (see section 2.2.2) actually integrates features of both the dual-route processing architecture and the polarity correspondence account. On the representational level, the dimensional overlap model assumes that dimensional overlap occurs when stimulus and response sets can be grouped into similar categories, which basically resembles the idea of polarity coding. On the processing level, the dimensional overlap model contains an unconditional, automatic route that programs an automatic response, as well as an intentional route that identifies the correct response according to task instructions, which is the same in the dual-route architecture assumed for SNARC. The only difference is, that for instance for the SNARC effect, the dimensional overlap model does not require the assumption that the unconditional response is programmed according to the spatial position of a number on the MNL.

In view of this, it appears that both the principle of polarity correspondence and the dimensional overlap model are strong frameworks to account for SNARC and other SRC effects. While polarity correspondence is a rather simple principle that can be applied to a variety of tasks and settings, the dimensional overlap model is more complex, which in turn allows for more precise predictions considering expected outcomes. Therefore, both these accounts will be revisited later for the question of independence of simultaneous SRC effects and the final discussion of collected data.

2.3.2 Mental Representation of Pitch

> *"How often do you hear a single sound by itself? Only when doing psychoacoustic experiments in a soundproof booth!"*
>
> (Christopher J. Darwin)

Sounds are an inextricable part of our environment at all times. An important attribute of sound is pitch, because it helps with the differentiation as well as with the grouping of acoustic information in our environment and is therefore "of primary importance in defining and differentiating our acoustic environment" (Plack & Oxenham, 2005, p. 3). Pitch is, for instance, important for speech communication, as it can carry prosodic (e.g., English), but also semantic information (e.g., Chinese). Furthermore, pitch and pitch perception are also central to music, as simultaneous and sequential presentation of tones forms the basis of harmony and melody (Schwartz & Purves, 2004).

Despite the criticism by Darwin (2005), who argued that in our everyday environment we usually perceive more than one sound at once, most research on pitch and pitch perception has focused and continues to focus on the perception of isolated tones. A reason for doing so is that some of the earlier research on SRC effects aimed to answer questions of applied psychology, like, for instance, how to best relate displays and controls. In line with this applied research, it appears more useful to study effects of isolated tones instead of a whole set of tones. Isolated tones are also used when representational organizations of pitch and pitch processing are investigated. Typically, early studies on the spatial representation of pitch required their participants to estimate the spatial location of the pitch height of a sound (e.g., Mudd, 1963; Pratt, 1930; Roffler & Butler, 1968; Stumpf, 1883; Trimble, 1934). It has only been more recently that spatial representations of pitch height have been investigated in a more indirect way, analogous to studies on the SNARC effect.

2.3.2.1 Defining Pitch and its Spatial Representation

Pitch is “the auditory attribute of a sound in terms of which sounds may be ordered on a scale extending from low to high. Pitch depends primarily on the frequency content of the sound stimulus, but it also depends on the sound pressure and the waveform of the stimulus” (ANSI, 1994). Subjectively perceived pitch is related to and primarily determined by the fundamental frequency of a tone (Moore, 2003; Schwartz & Purves, 2004). Even though it cannot be simply described as a function of frequency or any other physical stimulus parameter (Hall & Peters, 1981), in this work, pitch height will be referred to in terms of fundamental frequency (in Hz).

In the Western world, pitches are usually described as *low* or *high* (e.g., Casasanto, Phillips, & Boroditsky, 2003; Rusconi et al., 2006; Shayan, Ozturk, & Sicoli, 2011), which may even to a part form its vertical mental representation (Dolscheid, Shayan, & Casasanto, 2011). The question of whether this reflects actual spatial associations for pitch height was already assessed by Stumpf in 1883, who argued that the use of the vertical spatial dimension for describing pitch height was only metaphoric and that the auditory sensation did not include a real spatial characterization.

Contrary to the conclusion drawn by Stumpf (1883), later studies encouraged the idea of spatial representations of pitch height. Pratt (1930), who asked participants to locate the positions of tones on a numbered scale, observed that tones were located ascending on this scale according to their pitch height. He concluded that high tones seemed to be phenomenologically higher in space than low ones (Rusconi et al., 2006). Further studies produced similar results when participants were instructed to estimate the spatial position of sounds (e.g., Mudd, 1963; Roffler & Butler, 1968; Trimble, 1934; for a more detailed overview see Rusconi et al., 2006).

The structural relations between pitches can be described with geometrical models (Rusconi et al., 2006). Such models are derived from multidimensional scaling of pitch judgments (Shepard, 1982; Ueda & Ohgushi, 1987). Geometrical models describe pitch height in terms of an ascending spiral with circular (representing pitch chroma) and vertical rectilinear (representing pitch height) components (Ueda & Ohgushi, 1987). Therefore, as can be seen in Figure 4, geometrical models assume a more complex, two-dimensional, helix-shaped representation of pitch (Cohen Kadosh, Brodsky, Levin, & Henik, 2008a). This idea is supported by behavioral (Cohen Kadosh et al., 2008a), as well as neuropsychological (Warren, Uppenkamp, Patterson, & Griffiths, 2003) findings.

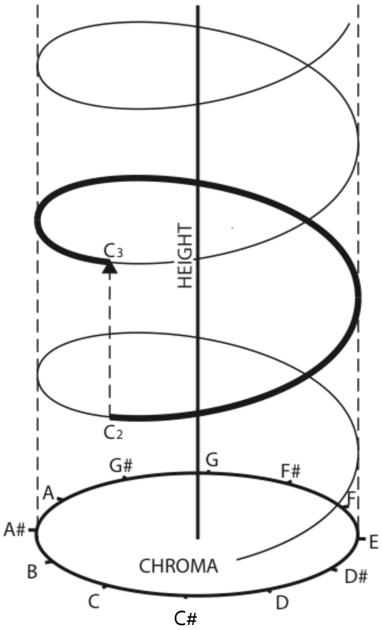

Figure 4: The pitch helix (Shepard, 1982)

Later studies on spatial associations of pitch height with representational space (e.g., Cho et al., 2012; Lidji et al., 2007) have confirmed a vertical rather than a horizontal orientation of this representation. In analogy with the MNL for numbers, Lidji et al. (2007) termed the ascending continuum, on which pitches are ordered from low to high, the *mental pitch line*.

2.3.2.2 Spatial Pitch Associations of Response Codes: The SPARC Effect

The pitch height of a sound is typically coded as a magnitude in choice reaction tasks, even when it is task-irrelevant. In tasks which require the judgment of a stimulus feature other than pitch (e.g., instrumental timbre), responses to tones high in pitch are made faster and more accurately with an upper hand key while responses to low tones are made faster with the lower hand key. This effect is termed the *Spatial Musical* (or *Pitch*) *Associations of Response Codes* (SMARC, Rusconi, Kwan, Giordano, Umiltà, & Butterworth, 2005; Rusconi et al., 2006; SPARC, Lidji et al., 2007). In agreement with Lidji et al. (2007), who stated that the pitch of isolated tones (as used in the present work) could not necessarily be considered music, the acronym SPARC will be adopted in this thesis.

Even though the acronym for spatial associations of pitch with response codes was only first formed in analogy to the term SNARC in 2005 by Rusconi et al., results on SRC for pitch height were already reported earlier. For instance, in 1982, a study was conducted by Shepard, who found a compatibility effect for the pitch height of a stimulus and spoken responses that differed in whether the vowel sound was high or low (i.e., /iː/ = high vs. /uː/ = low). A study of Walker and Ehrenstein (2000), who employed dynamic auditory stimuli (i.e., there was a pitch increase or decrease within a stimulus), used behavioral measures that were more similar to those of Rusconi et al. and Lidji et al.: Here, participants had to either respond to the initial pitch of the stimulus (*low* vs. *high*) or the direction in which the pitch changed (*lower* vs. *higher*) with an upper or lower response key. Results showed compatibility effects for pitch change, and, more importantly, for the initial pitch in which a tone was presented, regardless of task. Later studies investigating the SPARC effect aimed to investigate effects of spatial mapping of pitch in an isolated way and, therefore, used non-dynamic tones in which pitch was kept constant within a stimulus for both explicit (e.g., pitch comparison) and implicit (e.g., timbre judgment) SPARC tasks.

The vertical SPARC effect can be observed when responses are aligned vertically, that is, when there is an upper and a lower response key. The effect is commonly attributed to a vertical mapping of pitch heights onto representational space (e.g., Cho et al., 2012) which reflects a vertically oriented mental pitch line (Lidji et al., 2007). Vertical SPARC is typically independent of task and musical experience of participants (e.g., Lidji et al., 2007; Rusconi et al., 2006). This means that even participants with little or no former musical training (i.e., *nonmusicians*) automatically map pitch onto vertical responses, even when the pitch height of heard tones is not task-relevant.

When responses have to be made along a horizontal response alignment (i.e., responses are given with a left and right response key), a horizontal SPARC effect can be observed. Such an effect expresses itself through enhanced performance when tones low in pitch have to be mapped to a left key response and tones high in pitch have to be mapped to a right key response. This mapping is affected only by the location of the response key, but not by the responding hand (Rusconi et al., 2006). Horizontal SPARC effects are attributed to an orthogonal transformation of the original vertical pitch-to-space mapping into the horizontal dimension (Lidji et al., 2007; Nishimura & Yokosawa, 2009; Rusconi et al., 2006), which will be explained in more detail in the following section (2.3.2.3).

Horizontal pitch-to-space associations (i.e., horizontal SPARC) are known to be modulated by musical experience of participants (Cho et al., 2012; Lidji et al., 2007; Rusconi et al., 2006), task relevance of pitch (Cho et al., 2012; Lidji et al., 2007; Rusconi et al., 2006), and the presence of a reference tone (Cho et al., 2012). Generally, the remapping of the vertical representation onto the horizontal response alignment is less stable in nonmusicians. When the task is *explicit* (i.e., pitch is task relevant; e.g., pitch comparison), nonmusicians map pitch onto space automatically; when, however, the task is *implicit* (i.e., pitch his not task relevant; e.g., timbre judgment), results are inconsistent (SPARC observed: e.g., Nishimura & Yokosawa, 2009; Wolf, Bittrich, & Blankenberger, 2012; no SPARC observed: e.g., Lidji et al., 2007; Rusconi et al., 2006). A recent study suggests that within the view of polarity correspondence as an underlying mechanism, the relative coding of pitch as low or high is enhanced when a reference tone precedes the target (*referential coding*, Cho et al., 2012).

2.3.2.3 Mechanisms Underlying SPARC

Generally, vertical SPARC is assumed to reflect the mental representation of pitch along a vertically oriented continuum. Rusconi et al. (2006) concluded “that a representational dimension (pitch height) influences performance with vertically aligned responses irrespective of its relevance to the task. This suggests that our cognitive system maps pitch onto a mental representation of space” (Rusconi et al., 2006, p. 126).

In analogy with the assumptions made earlier for SNARC (see section 2.3.1.3), SPARC can be explained with the concept of a mental pitch line. Even though the spatial representation of pitch is assumed to be two-dimensional (e.g., Shepard, 1982), the primary dimension of this representation is the vertical axis which represents increasing frequency in terms of a mental line. Within a dual-route processing architecture, this mental pitch line causes unconditional responses to be prepared in accordance with a heard tone being low or high in pitch. However, like argued for the case of SNARC, SPARC has been observed also in the horizontal dimension, which questions the idea of a vertical pitch line to account for all spatial associations of pitch with responses. Like for SNARC, the dimensional overlap model (Kornblum et al., 1990), as well as the principle of polarity correspondence (Proctor & Cho, 2006), account for a wider range of pitch-to-space associations. For instance, according to the principle of polarity correspondence, high pitches, as well as upper or right key responses, are coded as + polarity. Thus, when a stimulus has to be mapped to a response with identical polarity, performance is enhanced.

The vertical spatial representation of pitch is thought to be the underlying basis for the observation of horizontal SPARC effects. As mentioned before, it is likely that horizontal SPARC effects result of an orthogonal transformation of the vertical pitch-to-space mapping into the horizontal response dimension (e.g., Rusconi et al., 2006). This transformation occurs as a consequence of the strong association of the categories *up* and *right* and the categories *left* and *low* (e.g., Cho & Proctor, 2003; Weeks, Proctor, & Beyak, 1995). In terms of polarity correspondence, this means that the response categories within the horizontal and vertical response set are coded as - and + polarity and fortuitous mapping leads to better performance.

According to Cho et al. (2012), the remapping of the vertical spatial representation of pitch into the horizontal dimension is facilitated in nonmusicians when they are provided with a task-irrelevant reference tone. This means that when such a reference tone with a constant, intermediate pitch precedes the target stimulus, participants unconditionally compare pitch heights. This comparison leads to the coding of the target pitch as (relatively) low or high. This relation is then coded as - or + polarity which in turn enhances performance for responses of the same polarity as the target.

Alternatively, the dimensional overlap model (Kornblum et al., 1990) would state that within a stimulus set, pitch heights, just like numbers, can be coded into categories. Also within this model, the presence of a reference tone would, as assumed by Cho et al. (2012), increase the likelihood of sounds being grouped into pitch categories which then leads to dimensional overlap between the stimulus feature pitch and the response set. However, dimensional overlap of pitch with horizontal responses is likely less strong than the overlap between pitch heights and vertically aligned responses, as pitch and vertical space share the same relation (low/high), while the relations of pitch and horizontal space are only similar (low/high and left/right).

In sum, the principle of polarity correspondence (Proctor & Cho, 2006) enables to explain both horizontal and vertical pitch-to-space associations (Cho et al., 2012) in a simplistic fashion, while, again, the dimensional overlap model (Kornblum et al., 1990) allows for a prediction of outcomes in greater detail. However, it is still unclear whether the observation that only nonmusicians show SPARC independently of experimental conditions can be solely attributed to the fact that, in nonmusicians, associations of pitch to vertical space may not be as overlearned as in musicians (e.g., Lidji et al., 2007). In order to further clarify what factors affect horizontal SPARC in nonmusicians, the present thesis works with SPARC as central concept of interest. Particularly, it is investigated, whether pitch height causes automatic horizontal spatial associations independently of other, accompanying stimulus features.

2.4 Simultaneous Stimulus-Response Compatibility Effects

Up to this point, the present thesis has brought forward details and accounts for SRC effects concerning isolated stimulus features with the examples of numerical magnitude and pitch. However, natural objects in our environment, as well as stimuli employed in experimental tasks, usually contain information on more than one stimulus dimension. When such stimuli are perceived, an integration of multiple stimulus codes is required which can lead to interferences in stimulus processing and response selection. In view of the concept of SRC it is of interest how such processing interferences can affect SRC effects of multiple stimulus features. In order to further investigate this question, this section begins with a view on the common processing of simultaneously perceived stimulus features. Next, the dimensional overlap model (Kornblum et al., 1990) and the principle of polarity correspondence (Proctor & Cho, 2006) are presented as frameworks that can account for simultaneous SRC effects. The section then continues with the question of whether horizontal SPARC in nonmusicians is independent of other auditory stimulus features and concludes with a more detailed look on timbre as a categorical and number as an ordinal variable that might interact with pitch.

2.4.1 Processing of Simultaneously Perceived Stimulus Features

When a stimulus with multiple features is perceived, all of these features are encoded and activated (Kahnemann & Henik, 1981; Kahnemann & Treisman, 1984; Tzelgov et al., 1992). This assumption entails potential interference between any kind of stimulus feature that is varied within a stimulus set. Whether multiple stimulus features interfere with each other or not depends on the way in which a stimulus is processed. In the past, there has been a long ongoing debate on whether multidimensional stimuli are processed holistically or analytically that cannot be answered in an all-or-none way (Kemler Nelson, 1993). Rather, since the 1960s, it is assumed that the type of processing depends on the type of stimulus (e.g., Garner, 1974; Lockhead, 1966; Lockhead, 1972; Shepard, 1964).

Stimuli within a set typically entail different features or dimensions on which they can vary and these dimensions can be classified as *integral dimensions* or *separable dimensions* (Garner, 1974). Garner (1974) introduced a set of four converging operations by which integral dimensions can be defined. Of interest for the present thesis are two of these operations, as they are related to speeded classification. One operation states that dimensional integrality is present when interference is observed in a speeded filtering task requiring selective attention to one dimension while the other one is varied. A very simple example for this kind of observation is a variant of the Stroop task (Stroop, 1935). In this task, participants are required to respond to the color in which a color word is printed. Thus, selective attention has to be paid to the color dimension (i.e., the color in which the word is printed in), while the other dimension (i.e., the color word) varies but has to be ignored. In this task it is typically observed that performance drops when print color and color word do not match. Therefore, visual color information and written color information can be considered integral dimensions. The second operation of interest introduced by Garner (1974) states that dimensions are integral when a correlation of dimensions leads to redundancy gain. In the Stroop example this would mean that in one block, participants would be presented with stimuli in which print color and color word are always consistent (i.e., the dimensions are positively correlated or congruent), while in the other block, the two dimensions would always be inconsistent (i.e., the dimensions are negatively correlated or incongruent). Typically, when dimensions are consistent, items are responded to faster than when they are inconsistent (i.e., there is a redundancy gain). When stimulus dimensions are not integral but separable, no such interference effects or redundancy gains are observed.

Garner's (1974) distinction of integral and separable dimensions can be re-formulated in terms of dependence and independence as investigated in the present thesis. As a consequence of the converging operations, integral dimensions are processed dependently, while the processing of separable dimensions is independent. This distinction is also valid for stimulus features that cause SRC.

In SRC tasks, stimulus dimensions or features can be grouped into three types: features that entail categorical information (e.g., number parity), features that entail spatial associations (e.g., numerical size), and features that are intrinsically spatial (e.g., stimulus location). Technically, any two of these three types can be varied along with each other in experimental tasks. Table 1 displays a matrix in which examples for each type of combination are presented. Note that the given examples only provide an illustration for possible feature type combinations; there is no statement made on whether the features or dimensions in these examples are integral or not.

Table 1: Illustrating examples for combinations of different types of simultaneously perceived stimulus features.

	categorical	**spatial association**	**intrinsically spatial**
categorical	color words printed in varying colors	numbers printed in varying colors	shapes presented on left or right side of a screen
spatial association		numbers printed in different sizes	numbers presented on left or right side of a screen
intrinsically spatial			words 'left' and 'right' presented on left or right side of a screen

2.4.2 Accounting for Simultaneous Stimulus-Response Compatibility Effects

When two stimulus features within a set can cause SRC in a given task, this means that both features overlap with the response. However, even though the overlap with the response could be seen as a conceptual similarity between the two stimulus features, this does not imply that these features are processed as integral dimensions. Whether two stimulus features that are consistent with responses are processed independently or not also depends on the degree to which they contain spatial information and whether the categories that can be formed for each feature contain similar relations (Henik & Tzelgov, 1982). Thus, SRC effects can jointly occur but still be independent.

In order to shed some light on possible underlying mechanisms for simultaneous SRC effects, in the following sections, the principle of polarity correspondence (Proctor & Cho, 2006), as well as the dimensional overlap model (Kornblum et al., 1990) will be revisited. In both cases, effects of integral and separable dimensions will be discussed.

2.4.2.1 Polarity correspondence of multiple stimulus features and responses

Within the framework of polarity correspondence (Proctor & Cho, 2006) it is stated that response selection will be faster when polarity codes of stimulus and response alternatives correspond (also see section 2.2.3). For the example of numbers brought forward earlier, this would mean that large numbers are responded to faster with a right hand response because both are coded as + polarity. However, the principle of polarity correspondence can also account for consistency effects other than isolated SRC.

When two features of the same stimulus can be coded as polarities, this can lead to enhanced performance when the polarity codes of these stimulus features correspond, resulting in a so-called *congruency* effect. For instance, a small number printed in a small font is congruent, because numerical and physical magnitudes in this example are both coded as the – polarity. In terms of Garner's (1974) distinction, a congruency effect resembles integral dimensions. It is important to note that such a congruency effect is only bound to the consistency of stimulus feature polarities, but not to that of responses.

When stimulus features are not integral but separable, no interference effects or redundancy gains will be observed (Kemler Nelson, 1993). However, it is still possible that both SRC effects present themselves independently of each other. Such an observation would mean that the polarity codes of both stimulus features and of the response set affect the response selection process.

2.4.2.2 Dimensional overlap between multiple stimulus features and responses

The dimensional overlap model (Kornblum et al., 1990) assumes an automatic route that activates an unconditional response when dimensional overlap between a stimulus feature and the response set is present (for more detail, see section 2.2.2). When the unconditionally programmed response is not consistent with the response that has been identified according to task instructions, the response time will increase because the unconditional response has to be aborted and the identified response has to be executed. However, the dimensional overlap model can also account for effects caused by more complex configurations of dimensional overlap between stimulus features and the response set. Table 2 is adapted from Kornblum and Lee (1995) and displays the eight possible types of stimulus-response ensembles that exist according to the taxonomy of the dimensional overlap model (Kornblum et al., 1990).

Table 2: Overview of stimulus-response ensembles according to the taxonomy of the dimensional overlap model (adapted from Kornblum & Lee, 1995)

	Overlapping ensemble dimensions			Illustrative stimulus and response sets		
	S-R dimensions		S-S dimensions	Illustrative stimulus sets		
Ensemble Type	Relevant (S_r)	Irrelevant (S_i)	(S_r-S_i)	Relevant (S_r)	Irrelevant (S_i)	Illustrative response sets
1	no	no	no	colors	geometric shapes	digit names
2	yes	no	no	digits	colors	digit names
3	no	yes	no	colors	digits	digit names
4	no	no	yes	colors	color words	digit names
5	yes	yes	no	colors	position (left or right)	keypresses (left or right) on colored keys
6	yes	no	yes	position (left or right)	colors and color words	keypress (left or right)
7	no	yes	yes	colors	color words/ position (left or right)	keypress (left or right)
8	yes	yes	yes	colors	color words	color names

The mechanisms that describe the experimental outcomes for each of these ensemble types are theoretically explained by the dimensional overlap model as presented by Kornblum et al. in 1990. However, a clearer and more detailed view is provided by the slightly expanded model presented in a paper by Kornblum and Lee in 1995 (Figure 5).

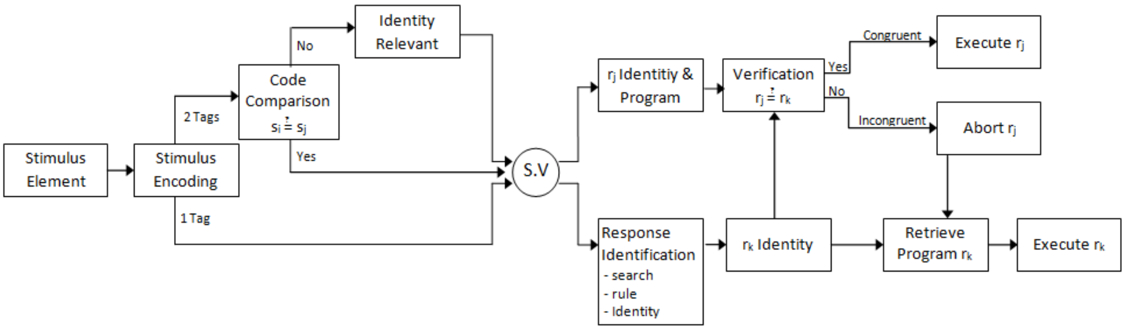

Figure 5: The extended dimensional overlap model (Kornblum & Lee, 1995)

In this version of the model, after the encoding of the stimulus, there are two different paths, depending on the amount of stimulus features that cause overlap (i.e., tags). When there is only one tag present, stimulus processing continues like in the original version of the model. When there are two tags present, the codes for these two stimulus features are compared for dimensional overlap. If they are the same (i.e., when there is dimensional overlap between S_i and S_j and stimulus codes are consistent or congruent, respectively), then this information determines the stimulus vector (S.V.) which then in turn enters the processing stage and activates the unconditional, automatic response. When two tags are present and the stimulus codes are not the same (i.e., there is dimensional overlap between S_i and S_j and codes are not consistent or incongruent, respectively), it is clear that for giving a correct response, the identity of this correct response (r_k), as determined by task instructions, has to be produced and the response program has to be retrieved. This, in turn, leads to a further increase of time until the response can be executed.

The present thesis aims to investigate whether horizontal SPARC in nonmusicians is independent of other, simultaneously perceived stimulus features, namely timbre and number. In a horizontal explicit SPARC task, participants respond to the pitch height of a heard sound. Therefore, the relevant stimulus dimension (S_r) entails dimensional overlap with the response. Number as a task-irrelevant feature (S_i) clearly also overlaps with the response set. Thus, stimuli containing numerical and pitch information may belong to the group of type 5 or type 8 ensembles, depending on whether pitch and number also share dimensional overlap (S_r-S_i). If such stimuli form a type 5 ensemble, the there will be no redundancy gain or interference effects between the pitch and number dimension (i.e., both SRC effects may be present but if so, they will only be additive). If this is not the case, the stimuli form a type 8 ensemble, where then congruency effects should be obtained.

For timbre, it is not entirely clear at this point whether it produces SRC effects or not. If it does, stimulus sets containing pitch as relevant (S_r) and timbre as irrelevant (S_i) dimension are type 5 or type 8 ensembles. If timbre does not produce any SRC effects, and therefore no overlap with responses, stimulus sets varying on the pitch and timbre dimension are type 2 or type 6 ensembles, depending on whether pitch and timbre share dimensional overlap. For a type 2 ensemble, a classic SPARC effect is the expected outcome. For a type 5 ensemble, the outcome may be a SPARC effect. However, it may also occur that the programming of an automatic response according to pitch is impaired due to interferences of pitch and timbre on the stimulus level.

In an implicit SPARC task, pitch height is a task-irrelevant feature. This means that there is dimensional overlap between the task-irrelevant dimension (S_i) and the response set. When pitch is combined with number as a relevant dimension (S_r), this again leads to type 5 or type 8 ensembles. The same is the case if pitch is combined with timbre under the precondition that timbre produces SRC as well. If timbre does not produce any SRC, this would mean that the task-relevant dimension (S_r) does not produce any dimensional overlap with the response set. Such stimulus sets are either type 3 or type 7 ensembles, depending on the relation between pitch and timbre. If there is no dimensional overlap between pitch and timbre, a standard implicit SPARC effect should be obtained. If there is dimensional overlap, such an implicit SPARC effect may not be observed.

Whether or not pitch and timbre or pitch and number cause dimensional overlap and are thus integral dimensions when varied within one and the same stimulus set, can only be clarified through experimental studies. In the following sections, outcomes on past studies related to this matter will be presented.

2.4.3 SPARC and Features of Auditory Stimuli: Dependent or Independent?

The perception of any kind of stimulus typically takes place on multiple dimensions. For instance, for a heard sound, not only its pitch, but also other features such as timbre, loudness, or duration, can be identified (e.g., Roederer, 2008). The principle of polarity correspondence states that polarity coding can take place for any kind of stimulus (feature) or response set, if the categories within a set can be coded asymmetrically (Proctor & Cho, 2006). Thus, a sound contains multiple kinds of information that can potentially be coded as polarities. This is for instance especially the case for duration and loudness, as these features are considered magnitudes as well. In addition, if a stimulus set contains more than one timbre, this feature can potentially be coded as polarities as well.

Considering the results of past studies (e.g., Lidji et al., 2007; Rusconi et al., 2006), it appears to be the case that vertical SPARC is generally stable, which implies that in the vertical response condition, mainly the polarity codes for pitch height affect responses. This is not surprising in view of the knowledge that most SRC effects, including spatial associations for duration and loudness, originally emerge in a horizontal response dimension, which may reduce interferences. Alternatively, it can be argued that in the vertical response condition, dimensional overlap of stimuli with responses is induced and reinforced due to the semantic categorization of pitches as *low* or *high*.

In the horizontal response dimension, it has been shown that pitch-to-space associations are less stable in nonmusicians (Cho et al., 2012; Lidji et al., 2007; Rusconi et al., 2006). When responses are made with horizontally aligned keys, polarity codes of other stimulus features that are usually represented along or associated with the horizontal axis might modulate polarity correspondence of pitch and response or, in terms of dimensional overlap, the stimulus feature that causes the unconditional, automatic response. Therefore, it is important to consider whether horizontal SPARC emerges independently of other stimulus features or not. In particular, it is of interest whether possible interferences emerge only for stimulus features that can be considered magnitudes, like numerical magnitude, or also for stimulus features that only contain structural similarity with the response set, like timbre.

2.4.3.1 SPARC and Timbre

Previous research on the horizontal SPARC effect has identified the factors task (explicit vs. implicit), participant subgroup (musicians vs. nonmusicians), and procedure (reference tone vs. no reference tone) to affect the mapping of pitch onto horizontal space (e.g., Cho et al., 2012; Lidji et al., 2007; Rusconi et al., 2006; for more details, see section 2.3.2.2). It is important, however, to also take properties of the stimulus set into consideration. When any kind of structural similarity is sufficient to produce overlapping polarity codes of stimulus and response, participants might as well code stimulus attributes other than the pitch of a heard sound as polarities (e.g., timbre). This, in turn, might then interfere with response selection based on pitch polarity and affect participants' response behavior.

Timbre is, next to pitch and loudness, one of the three primary dimensions of sound (Melara & Marks, 1990a; Melara & Marks, 1990b). According to Levitin (1999), it is the most important one of the six perceptual attributes of auditory events. Timbre is the auditory feature that allows to distinguish between different types of sound production (i.e., the identity of a sound source; e.g., Caclin, Giard, Smith, & McAdams, 2007), even when the other features remain the same and it is mainly determined by the specific quality of the frequency spectrum (e.g., Roederer, 2008). Even though timbre is often defined in a categorical fashion, humans are able to distinguish timbres along different dimensions such as for example attack time, spectral centroid, and spectral flux (McAdams, Winsberg, Donnadieu, De Soete, & Krimphoff, 1995; Menon et al., 2002) and it has been sown that such timbre dimensions are processed in an interactive fashion (Caclin et al., 2007). Despite this, in this work, timbre will be considered in terms of a categorical variable. The reason for doing so is that in the experiments investigating relations of SRC effects of pitch and timbre in the present thesis, only two timbres that are very easily distinguishable are employed.

In the stimulus sets of previous SPARC studies, timbre has been varied to a great extent (single instruments: e.g., piano, violin, clarinet; instrument groups: e.g., percussions, wind instruments, brass instruments, string instruments; non-instrumental sounds: e.g., sinusoids), but has often not been included as a factor in analyses. Timbre and pitch are, however, not perceived independently (e.g., Beal, 1985; Crowder, 1989; Krumhansl & Iverson, 1992; Melara & Marks, 1990a; Melara & Marks, 1990b; Russo & Thompson, 2005; Singh, 1992; Vanzella, Schellenberg, & Bishop, 2010; Warrier & Zatorre, 2002; Zarate, Ritson, & Poeppel, 2013). Participants cannot ignore the task-irrelevant variation of one of these two attributes while attending to the other, task-relevant one (Pitt, 1994). Moreover, especially for nonmusicians it is difficult to ignore timbre changes when judging pitch and vice versa (Krumhansl & Iverson, 1992). This is particularly interesting because it means that pitch and timbre can share dimensional overlap. In addition, for nonmusicians, timbre is a more salient sound attribute and weighed more heavily than pitch (Pitt, 1994; Wolpert, 1990). Thus, even when timbre categories are easily distinguishable, they have the potential to affect pitch-to-space associations. Therefore, timbre has to be considered as a possible modulating factor on automatic mappings of pitch to space.

The relative pitch height of a sound within a given set and its tonal distance from the middle of this set is yet another important attribute by which sounds can be classified. Tonal distance can affect performance in speeded choice-reaction tasks such as an explicit or implicit SPARC task. The greater the distance between two pitches, the easier they are discriminable (e.g., Cohen Kadosh et al., 2008a; Elkin & Leuthold, 2011). Pitch distances have previously been shown to interact with pitch-to-space associations. For instance, for relatively great pitch distances, the strength of the observed SPARC effect increases (Beecham, Reeve, Wilson, & Antonietti, 2009) while for relatively small distances, the effect may disappear. When tonal distances are within the range of half an octave, SPARC can even be reversed (Rusconi et al., 2006). In sum, the pitch distances within a stimulus set and the resulting tonal range potentially affect the spatial mapping of pitch and interferences with other stimulus features to the extent that possible effects are concealed if pitch distances are not considered as a factor in analyses.

To summarize, prior studies on the automatic mapping of pitch height to space have shown a dependency of musical experience, response alignment, and the presence of a reference tone (e.g., Cho et al., 2012), while little attention has been paid to sound attributes other than pitch, such as timbre and pitch distance or pitch range, respectively. These factors, however, have the potential to modulate automatic spatial mappings of pitch heights. This might be particularly relevant for nonmusicians, which have been shown to produce less stable automatic pitch-to-space mappings.

The aim of Study 1 in the present thesis is therefore to investigate the influence of timbre on automatic coding of pitches as low and high in nonmusicians. Further details on experimental variation and methods are given in section 3.1.

2.4.3.2 SPARC and SNARC

The principle of polarity correspondence (Proctor & Cho, 2006; sections 2.2.3 and 2.4.2.1) and the dimensional overlap model (Kornblum et al., 1990; sections 2.2.2 and 2.4.2.2) enable to account for effects that may be observed when polarity codes for features of one and the same stimulus correspond or, in other words, when there is dimensional overlap between multiple stimulus features. In particular, it allows for explaining possible interactions of SPARC with other stimulus features that have not been considered in the past. Specifically, also interferences of SPARC with representations of other magnitudes and resulting effects can be accounted for.

Timbre is a stimulus feature that can, at least in the studies of the present thesis, at the most produce structural overlap with a response set, due to its categorical nature. In order to gain a more complete picture, however, it is of interest to also investigate whether pitch-to-space associations, which are thought to originate from a vertical spatial representation, are affected when varied along with another magnitude representation that originates in the horizontal dimension. Numerical magnitude is a very suitable stimulus feature for this, as spatial-numerical associations have been shown to be relatively persistent (Cohen Kadosh et al., 2008c), specifically also in the auditory domain (e.g., Cohen Kadosh et al., 2008c; Nuerk et al., 2005).

When numerical magnitude is varied along with other magnitudes within a stimulus set, congruency effects can be observed. Such effects involving numbers are typically termed as *size congruity* effects (e.g., Henik & Tzelgov, 1982; Schwarz & Heinze, 1998). A prominent example for the size congruity paradigm is the combined variation of numerical and physical size of stimuli (e.g., Algom, Dekel, & Pansky, 1996; Cohen Kadosh & Henik, 2006; Henik & Tzelgov, 1982; Pansky & Algom, 1999; Santens & Verguts, 2011; Schwarz & Ischebeck, 2003). In this paradigm, participants are usually presented with two digits of different physical and numerical sizes and are asked to judge which digit is (either numerically or physically) larger by pressing the response key on the side on which the (numerically or physically) larger stimulus appeared. When stimuli are congruent, that is, when a stimulus is both physically and numerically larger (e.g., 2, 8), responses are faster than when physical and numerical size of the stimuli do not correspond (e.g., 2, 8). Size congruity effects can also be observed when only a single digit is presented (e.g., Santens & Verguts, 2011; Schwarz & Heinze, 1998; Schwarz & Ischebeck, 2003). In this kind of task, participants are asked to compare the presented stimulus to a given standard (e.g., whether its numerical magnitude is smaller or larger than 5).

Effects of size congruity have also been reported in studies that investigated interactions between numerical magnitude and duration (Dormal, Andres, & Pesenti, 2008; Dormal, Andres, & Pesenti, 2012), dot size (Gebuis, Cohen Kadosh, Haan, & Henik, 2009), luminance (Cohen Kadosh et al., 2008b), and line length (Dormal et al., 2012; Dormal & Pesenti, 2007). The observed interferences are generally attributed to be a consequence of shared magnitude processing (Cohen Kadosh et al., 2008c).

The exact processing stage on which interference due to size congruity is thought to take place has been further specified through an extensive study by Santens and Verguts (2011). These authors tested whether the interaction of numerical and physical size derived from the representational or rather from the decisional level of stimulus processing. While the *shared representations account* attributes observed interferences to the mapping of both numerical and physical magnitudes onto a shared representation, the *shared decisions account* states that this interaction takes place on the decision level (Schwarz & Heinze, 1998, see Figure 6).

In their study, Santens and Verguts (2011) found that congruity effects were task-dependent. Specifically, when the task required the comparison of (numerical or physical) size, congruency effects were observed, whereas when the task involved the judgment of number parity, congruency did not facilitate responses. These results imply that congruency effects derive on a decisional level rather than on a representational one. This is because in Santens and Verguts' experiments, only the required decision, but not the stimulus set itself was varied. By implementing the shared decisions account into a computational model, Santens and Verguts were able to simulate congruency effects analogous to the behavioral results in all their experimental tasks, confirming the assumptions based on the shared decisions account.

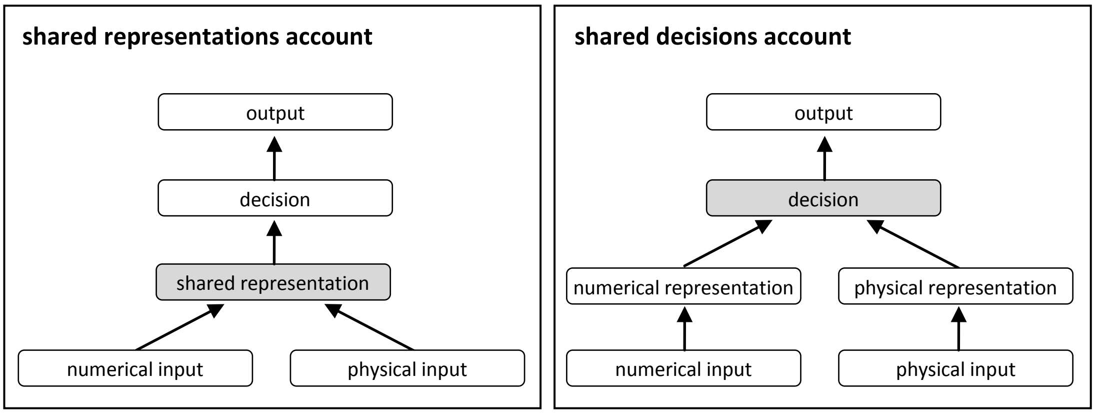

Figure 6: Models for the shared representation account (left panel) and the shared decisions account (right panel) for congruity effects of numerical and physical size. The grey field depicts the origin of the size congruity effect. (Santens & Verguts, 2011)

The research of Santens and Verguts (2011) shows that for the case of numerical and physical size in the visual domain, the interference effect in holistic stimulus processing takes place on a decisional level. Based on their findings, the authors state that the dual-route model they present in their work predicts that novel size congruity effects of numerical size and other quantities also originate on the decision level of stimulus processing. In particular, Santens and Verguts predict that "each quantitative dimension can interfere with any other to the extent that the irrelevant dimension is effective in automatically activating the decision units that are used for the task at hand" (Santens & Verguts, 2011, p. 107). This prediction is further supported by the results of several neuroimaging studies which point towards a shared neural code for magnitude processing (e.g., Cohen Kadosh et al., 2005; Cohen Kadosh et al., 2008c; Fias, Lammertyn, Reynvoet, Dupont, & Orban, 2003; Vogel, Grabner, Schneider, Siegler, & Ansari, 2013) that is independent of the input format. Specifically, Cohen Kadosh et al. (2008c) state in their review, that the collected results "suggest that the IPS [i.e., intraparietal sulcus] hosts overlapping domain-general and domain-specific neural populations in human adults for numbers and different magnitudes" (Cohen Kadosh et al., 2008c, p. 140). Thus, when including the variation of numerical size and pitch within one task, these magnitudes may also interfere on a behavioral level depending on the decision required for the task.

In view of this, it is of interest whether SNARC and SPARC, which originate from different representational dimensions, are independent. This possibility is suggested by a study in which participants completed both an explicit and an implicit task for both SNARC and SPARC in a blocked design (Beecham et al., 2009). For each task, the data were submitted to a hierarchical cluster analysis in order to identify whether participants showed a SNARC or SPARC pattern, a reversed SNARC or SPARC pattern, or none of these. The authors found no significant relationship between subgroup memberships, which implies that SNARC and SPARC were not related. However, the method of Beecham et al. (2009) is a rather indirect way to investigate overlap between SNARC and SPARC, as in this study, number and pitch height were varied in separate tasks. A factorial design, in which both SNARC and SPARC compatibility are varied within one and the same task, is a more direct way to answer the question of independence.

An optimal way for implementing such a factorial design was found by Fischer et al. (2013). These authors combined numerical magnitude and pitch height within auditorily presented stimuli (number words ranging from two to six sung in five different pitch heights). While in half of the blocks their nonmusician participants decided whether a sung number was smaller or larger than the immediately preceding reference (always "four" sung in intermediate pitch), in the other half they judged whether the pitch of the target sound was lower or higher than the reference. Responses were given along a diagonally aligned response set containing a lower left and an upper right response key. The central finding was that, in both tasks, only the task-relevant magnitude produced an SRC effect. In other words, instead of polarity categories of both stimulus attributes, only the polarity of the relevant attribute affects RT. According to Fischer et al., these results imply that numerical magnitude and pitch are not automatically mapped onto spatial representations. Rather, participants employed a cognitive strategy that selectively maps only the task-relevant attribute to its specific spatial dimension (the Strategic Association of Response Codes, STARC hypothesis; Fischer, 2006).

The results of Fischer et al. (2013) imply that SNARC and SPARC do not interact. Specifically, no congruency effects were obtained, even though the study used a diagonal response set which should, according to the authors, have allowed for capturing both SNARC and SPARC at the same time. As mentioned earlier, congruency means that two stimulus attributes share dimensional overlap or that they are coded as the same polarity. For instance, “nine” sung in a high pitch is congruent, because numerical magnitude and pitch height have the same polarity, whereas “nine” sung in a low pitch is incongruent (e.g., Boenke, Ohl, Nikolaev, Lachmann, & van Leeuwen, 2009). Such congruency effects are typically considered to result from automatic, preattentive binding across stimulus dimensions (Pomerantz & Lockhead, 1991). The absence of these effects in the study of Fischer et al. is therefore consistent with the authors’ claim that magnitude-to-space associations are not produced automatically.

Nevertheless, the results of Fischer et al. (2013) do not support these strong conclusions due to tasks and stimuli employed in their study. In their magnitude comparison tasks, one effect is always explicit while the other one is implicit. The absence of the implicit effects, however, does not necessarily imply that it is always only the task-relevant magnitude that is spatially mapped onto responses. Furthermore, implicit spatial associations of magnitudes can be influenced by the range from which specific magnitudes are chosen. While for numbers the explicit SNARC effect occurs for relatively small distances, in implicit SNARC tasks (i.e., where numerical magnitude is not relevant for task solving), this is not always the case, as the *distance effect* (Moyer & Landauer, 1967) leads to increased difficulty of number discrimination when numerical distances are very small (e.g., Dehaene, 2011; Verguts & van Opstal, 2005). Moreover, the number 5 has a special status on the internal number scale (Tzelgov et al., 1992), as it forms the midpoint of a decade, which is why within a range from 0 to 10 it is likely best to choose stimuli around the midpoint 5. As will be shown in Study 1, a small range is an even larger problem with implicit SPARC. Nonmusicians do not always automatically map pitch height onto space (e.g., Cho et al., 2012; Lidji et al., 2007; Rusconi et al., 2006), but the mapping process can be reinforced by the presence of a reference tone (Cho et al., 2012) and potentially by factors such as timbre or pitch range, as investigated in Study 1. In particular, the finding of an implicit horizontal SPARC effect in nonmusicians with a consistently presented reference tone (Cho et al., 2012) contrasts with the results of Fischer and colleagues.

In sum, when varying pitch height and numerical magnitude within one stimulus set, these magnitudes can potentially interact. Prior studies (Beecham et al., 2009; Fischer et al., 2013) investigating this question cannot terminally answer the question of independence of pitch and number processing due to methodological deficiencies. Therefore, the aim of Study 2 is to introduce the paradigm of Fischer et al. (2013) to a set of tasks and stimuli sufficient to produce a SNARC and SPARC effect also in implicit task conditions. Details on experimental methodology and possible outcomes are described in section 4.1.

2.5 Overview and Research Interest

Effects of SRC, like for pitch heights, can potentially be modulated by further stimulus features varied within a stimulus set. The horizontal SPARC effect in nonmusicians has been shown to be less stable than pitch-to-space associations in other participant subgroups (i.e., trained musicians) or in other response dimensions (i.e., the vertical response dimension). This may be because in the horizontal dimension, horizontally represented additional stimulus features can be consistent with responses.

In order to further clarify the inconsistency of findings on the horizontal SPARC in nonmusicians, the present thesis investigates how these pitch-to-space associations interact with further stimulus features. As lined out in section 2.4.1, there are three types of stimulus features: features that are of categorical nature (e.g., timbre), features that entail spatial associations (e.g., numbers) and features that are intrinsically spatial (e.g., presentation side of auditory stimuli). These features can produce SRC effects with horizontal responses originating from a horizontally aligned representational organization. The question of independent processing of pitch and intrinsically spatial features is not addressed in this thesis as such an experiment has already been carried out by Nishimura and Yokosawa (2009). The presentation of auditory stimuli that varied in pitch and the side of the ear to which they were presented yielded SRC effects of both pitch (i.e., a SPARC effect) and presentation side (i.e., a Simon effect; Simon & Rudell, 1967) that were independent. Therefore, in the present thesis, two studies, in which additional stimulus features entailing categorical information or spatial associations were varied, were conducted consecutively. Study 1 investigates the simultaneous processing of pitch and timbre under conditions of varied pitch range. Study 2 was conducted hereafter and investigates the question of independence of pitch and number processing. Results obtained in Study 1 were taken into account for the construction of the stimulus set in Study 2.

3 Study 1: Pitch and Timbre

3.1 Research Interest

The aim of Study 1 in this thesis is to investigate the influence of pitch and timbre differences on the automatic coding of pitches as *low* and *high* in nonmusicians. This implies a variation of pitch differences (i.e., pitch range) and response alignment while keeping constant timbre categories. A set of three timbre judgment tasks was conducted with piano and vocals as timbre categories. The first task involved sounds within a small pitch range which had to be responded to with horizontally aligned responses (Experiment B). The second task differed from the first one only in the orientation of the response alignment, which was vertical (Experiment C). The third task adopted the horizontal response alignment of the first task, but involved a set of stimuli with increased tonal distances (Experiment D). In order to assure that the specific set of stimuli was suitable to produce explicit (and therefore not necessarily automatic) pitch-to-space associations, in a prior control experiment, participants completed a pitch comparison task with the stimuli of the first timbre judgment task (Experiment A).

These experiments provide the opportunity to observe possible isolated effects of timbre and tonal distance as well as interactions of such effects with effects of pitch-to-space mappings. Particularly, it might be observed that pitch interferes with timbre judgment and vice versa, since, as argued before, pitch and timbre are not processed independently (e.g., Krumhansl & Iverson, 1992).

If interference is observed in a timbre judgment task with pitch height varying as a task-irrelevant feature, this would, within the taxonomy of the dimensional overlap model (Kornblum et al., 1990), mean that pitch and timbre form a type 7 or a type 8 ensemble, depending on whether timbre produces dimensional overlap with the response or not (see section 2.4.2.2 for more details). If no interference effects are observed, pitch and timbre form a type 3 or a type 5 ensemble. If interference is observed in a pitch comparison task with timbre varying as a task-irrelevant feature, pitch and timbre form a type 6 or a type 8 ensemble. If no such effects are observed, pitch and timbre form a type 2 or a type 5 ensemble.

In both tasks, the observation of interference effects would imply that pitch and timbre are indeed integral stimulus dimensions of auditory stimuli, which is indicated by the results of a range of studies which observed interaction between pitch and timbre on a perceptual level (e.g., Krumhansl & Iverson, 1992). The results will have also relevance for the more general question of whether mappings of pitch to horizontal space can be investigated without considering other stimulus attributes.

3.2 Empirical Section

In the empirical section of Study 1, each experiment is introduced with a short summary of theoretical facts that motivate the experiment (however, for a more detailed introduction to the general topic, see section 2.4.3.1), followed by methods and results. Also, short discussions and transitions between experiments are provided, whereas a joint discussion of results follows in section 3.3.

3.2.1 Experiment A (Control Task): Pitch Comparison in a Horizontal Task Setting With Small Pitch Range

Past studies on the horizontal SPARC effect in nonmusicians have reported consistent findings only when pitch height was task-relevant (e.g., Cho et al., 2012; Lidji et al., 2007; Rusconi et al., 2006). For solving such a task, pitch height has to be categorized as *low* or *high* (or *lower* or *higher* than a given standard). This induced structural asymmetry within the set of pitches leads to polarity coding and facilitates the mapping to left and right key responses.

As timbre and pitch of a sound are not perceived independently (e.g., Krumhansl & Iverson, 1992), it is an interesting question whether the judgment of timbre is affected by pitch and vice versa. In order to also determine whether timbre contributes to or impairs the horizontal SPARC effect in nonmusicians, a pitch comparison task was conducted in which half of the sounds were piano tones and the other half were vocal tones. These timbre categories were maintained throughout the whole set of experiments in Study 1 while varying other factors. This approach allows for comparing findings across tasks and for testing whether vocal sounds as a novel timbre category (in contrast to instrumental sounds and sinusoids) lead to similar pitch-to-space mappings as compared to those for piano tones and other instruments.

3.2.1.1 Method

Participants. Twenty-four students of the University of Kaiserslautern were paid for their participation (15 female; average age: 24.9 years; range: 21 – 30 years). All of them were right-handed and reported normal hearing. Fourteen of them had not received any musical training in the past; the others had received an average of 4.9 years of musical training, but had stopped since 9.1 years on average. Therefore, they were all considered as nonmusicians.

Stimuli and procedure. Stimuli consisted of a set of two low- and two high-pitch tones (A3, B3, B4, and C#4, respectively 220.00, 246.94, 493.88, and 554.37 Hz) with distances of 6 and 8 semitones around the intermediate reference (F4, 349.23 Hz) in two timbres, that is, piano and vocal tones. The vocals consisted of the syllable /ba:/ in German pronunciation, sung by a proficient female singer, recorded through a Sennheiser microphone (MD 421-II, Sennheiser electronic) and an Edirol external sound card (USB Audio Capture UA-25, Edirol) to a MacBook Pro (Mac OS X Version 10.6.3, Apple) via the Amadeus Pro (v1.5.1, HairerSoft) software. The recordings took place in a professional audio cabin and were edited with Audition 5 (Audition CS5.5, Adobe Systems Inc.). The piano sounds were produced via the Akoustik KeyZ (DSK Music) plug-in with Audition 3 (Audition 3.0, Adobe Systems Inc.). All stimuli were edited to a duration of 600 ms.

Participants were tested individually in a soundproof booth. The experiment was run on a Dell laptop computer (Latitude D380, Dell). Stimulus presentation and response collection were controlled through the E-Prime software (E-Prime 2.0, Psychology Software Tools Inc.). Stimuli were presented via closed Beyerdynamic headphones (DT-770 Pro, Beyerdynamic) at equal loudness well above threshold and clearly within the individual comfort zone.

On each trial, a fixation cross was presented in the centre of the laptop screen for a duration of 200 ms; 250 ms after the onset of the fixation (i.e., 50 ms after the fixation had disappeared), the reference tone was presented for a duration of 600 ms, followed by the target sound in the respective timbre after an inter-stimulus interval (ISI) of 500 ms. After each target, participants were required to judge whether the target was lower or higher in pitch than the reference by pressing one of two horizontally aligned response keys (Q or P) with their left or right index finger. Responses were recorded until 2500 ms after stimulus onset, followed by an inter-trial interval (ITI) of 1000 ms. The time course of a trial is displayed in Figure 7.

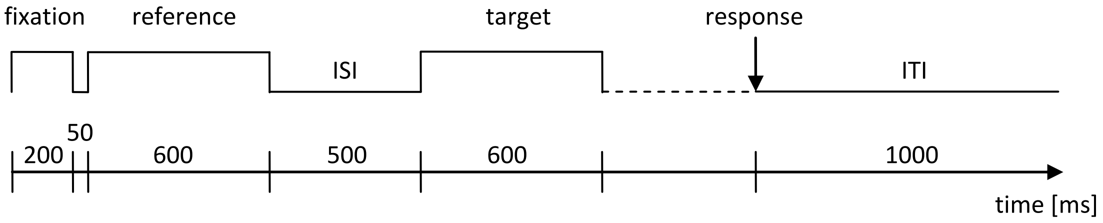

Figure 7: Time course of a trial in the pitch comparison task in Experiment A.

Participants completed two experimental blocks with 80 trials each (10 repeated measures per stimulus) that differed only in the assignment of the lower and higher category to response sides. Trials were presented in a pseudorandom order with the restriction that two consecutive trials did not contain the same target sound and that there were no more than three consecutive trials with the same correct response location. The order of blocks was counterbalanced across participants. Each block was preceded by a practice block consisting of 8 trials with feedback (RT in ms and correctness displayed on the screen after each trial). Altogether, the experiment lasted about 25 minutes.

3.2.1.2 Results

Reaction times shorter than 100 ms were excluded from analysis (3 of 3840 trials). In addition, an individual outlier criterion of two standard deviations was applied to the data of each participant (176 trials, 4.6%, approximately equally distributed across conditions). Mean differences (right – left key) in RT (dRT, correct trials only) and error rates (dER) are displayed in Table 3 as a function of timbre, global pitch, and distance from the middle.

Table 3: Average dRT (ms) and dER (%) in Experiments A, B, C, and D for each timbre, global pitch, and tonal distance from the middle.

	piano				vocals			
	low		high		low		high	
	far	close	close	far	far	close	close	far
Experiment A	pitch comparison horizontal – small range							
dRT (ms)	27.6	56.8	-70.6	-55.9	55.6	42.5	-22.4	-63.6
dER (%)	-0.8	7.1	-4.8	-6.7	2.1	1.3	-2.5	-2.1
Experiment B	timbre judgment horizontal – small range							
dRT (ms)	-4.2	8.6	-35.5	-41	11.8	2.0	-13.1	2.0
dER (%)	2.9	7.5	-7.3	-4.2	5.8	5.8	-2.9	-3.3
Experiment C	timbre judgment vertical – small range							
dRT (ms)	14.9	1.7	-17.6	-19.7	20.8	2.2	-30.4	-19.0
dER (%)	5.0	2.1	-2.7	-4.6	4.2	-2.5	-4.2	-2.5
Experiment D	timbre judgment horizontal – large range							
dRT (ms)	-21.8	3.4	-55.4	-43.3	24.2	10.0	-31.6	-12.3
dER (%)	3.8	7.9	-6.3	-11.3	10.0	6.3	-6.3	-4.6

Note. Horizontal = left and right response key; vertical = upper and lower response key; dRT = difference in mean RTs; dER = difference in ERs.

There was no correlation of RT and error rates (ER), $r = .083$, $p = .698$, indicating that there was no speed-accuracy tradeoff. A 2 × 2 × 2 analysis of variance with repeated measures (rmANOVA) with the within-subject variables Pitch (low vs. high), Timbre (piano vs. vocals), and Distance (6 vs. 8 semitones from reference) was conducted on the mean dRT and dER for each dimension. ER analyses were calculated on arcsine-transformed ER. Results of statistical analyses on dRT and dER are provided in Table 4. Note that for easier reading and interpretation of the Graphs, results for dER are displayed as untransformed values (in %).

Table 4: Results for the statistical analyses on RT and ER in Experiment A.

	F	p	η_p^2
RT			
Pitch	9.63	.005	.295
Timbre	5.84	.024	.203
Distance	4.42	.047	.161
Pitch × Timbre	1.43	.244	.058
Pitch × Distance	0.25	.621	.011
Timbre × Distance	0.41	.529	.017
Pitch × Timbre × Distance	14.23	.001	.382
ER			
Pitch	26.80	< .001	.538
Timbre	1.78	.195	.072
Distance	4.08	.055	.151
Pitch × Timbre	2.83	.106	.110
Pitch × Distance	0.76	.392	.032
Timbre × Distance	5.06	.034	.180
Pitch × Timbre × Distance	1.25	.276	.051

RT analyses. The ANOVA on RT revealed a main effect of Pitch, *F*(1, 23) = 9.63, *p* = .005, with a 45.6 ms advantage for the low-left mapping and a 53.1 ms advantage for the high-right mapping (see Figure 8). There was also a main effect of Timbre, *F*(1, 23) = 5.84, *p* = .024, meaning that piano sounds were on average responded to 13.6 ms faster with the right hand than vocals. In addition, a main effect of Distance was found, *F*(1, 23) = 4.42, *p* = .047, with greater distances on average being responded to 10.7 ms faster with the right hand. These main effects were further qualified by a three-way interaction of Pitch, Timbre and Distance, *F*(1, 23) = 14.23, *p* = .001.

Separate analyses for each timbre revealed a two-way interaction of pitch and distance both in the piano category, *F*(1, 23) = 5.95, *p* = .023, and the vocal category, *F*(1, 23) = 12.72, *p* = .002. However, the directions of these interactions were opposed to each other: While in the piano category, greater distances lead to smaller dRTs (29.2 ms) for high pitches, the opposite was found in the vocals category, where dRTs increased for greater distances in low pitches (41.3 ms). None of the other effects reached significance.

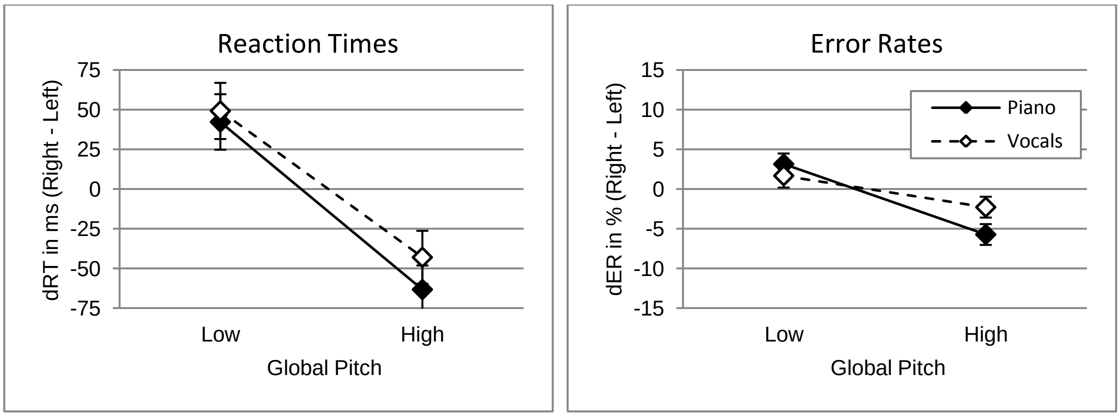

Figure 8: Mean dRT (left panel) and dER (right panel) for the pitch comparison task with small pitch range and horizontally aligned responses in Experiment A.

ER analyses. The analyses on arcsine transformed ER yielded a main effect of Pitch, *F*(1, 23) = 26.80, *p* < .001, with a 2.4% advantage for the compatible low mapping and a 4.0% advantage for the compatible high mapping. There was a two-way interaction of Timbre and Distance, *F*(1, 23) = 5.06, *p* = .034. Separate analyses revealed a main effect of Distance only in the piano category, *F*(1, 23) = 6.88, *p* = .015, where greater distances were on average responded to 4.9 ms faster with the right hand. None of the other effects reached significance.

3.2.1.3 Discussion

In line with previous studies that employed speeded bimanual pitch comparison tasks (Cho et al., 2012; Lidji et al., 2007; Rusconi et al., 2006), faster and more accurate responses were observed for the low/left- and right/high-mapping (i.e., a SPARC effect). Thus, performance was enhanced when polarity codes of pitch height and response overlapped.

In accordance with a recent study in which participants were asked to judge the pitch height of sung number words (Fischer et al., 2013), the observed SPARC effect was found to be present also in the vocal timbre category. Specifically, timbre did not modulate pitch-to-space mappings in the nonmusician participants when performing pitch comparisons. This similarity in findings for piano and vocal sounds is to be expected as vocal sounds are comparable to instrumental sounds or can even be considered as a special form of instrumental sounds due to the nature of their production and their line of harmonics (Roederer, 2008).

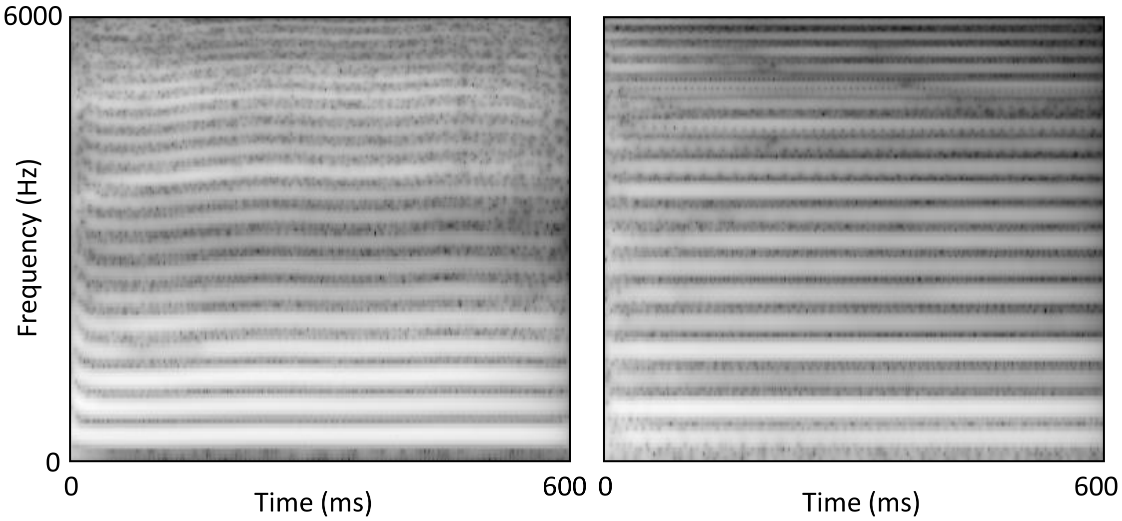

Figure 9: Spectrogram of the reference tone in the vocal category (left panel) and the piano category (right panel).

The similarity of the timbre categories used in the present experiment is illustrated in Figure 9. Here, the spectograms of the reference tones for both piano and vocals are depicted. Note that the formants (i.e., the bright horizontal frequency bands) for both timbres lie on the same frequencies, which means that the sounds are of the same pitch, but also that they contain the same line of harmonics.

The present results show that the set of stimuli that was used in Experiment A is suitable to produce pitch-to-space associations in a pitch comparison task. More specifically, explicit judgments of pitch height were not affected by timbre, even though timbre was associated with response sides. Therefore, pitch and timbre can be viewed as type 5 ensemble for dimensional overlap in this task. In order to test whether timbre modulates mappings of pitch height onto horizontal responses in an implicit task, it is interesting to compare the findings of Experiment A with results obtained in a horizontal timbre judgment task with nonmusicians.

3.2.2 Experiment B: Timbre Judgment in a Horizontal Task Setting With Small Pitch Range

The results of Experiment A showed that the SPARC effect in nonmusicians was not modulated by timbre category in a horizontal pitch comparison task. This finding goes in line with the assumption of polarity correspondence of left responses with low tones and right responses with high tones (Cho et al., 2012).

Experiment B aimed to investigate, whether the automatic mapping of pitch height onto horizontal responses in an implicit SPARC task (i.e., pitch is not task-relevant) can be affected by timbre. In this experiment, the stimulus and response set of Experiment A (pitch comparison task) were adopted for an implicit SPARC task in which participants had to respond to the timbre category of the tone (piano or vocals). Technically, based on the results of Experiment A, it would be expected that also in Experiment B, timbre and pitch form a type 5 ensemble and are therefore processed independently.

Previous studies did, however, not lead to consistent conclusions regarding the question whether horizontal pitch-to-space mappings in nonmusicians are generated automatically. This may be due to the fact that these studies did not sufficiently consider that pitch is not perceived independently from timbre (e.g., Pitt, 1994).

Therefore, when pitch is the irrelevant stimulus dimension in a speeded bimanual choice reaction task, dimensional overlap between pitch and timbre may conceal SPARC. The present experiment enables to observe possible interactions of timbre with the spatial mapping of pitches in nonmusicians. Such an interaction would imply that interferences of timbre and pitch are dependent on the task relevance of pitch for this group of participants.

3.2.2.1 Method

Participants. Twenty-four students (14 female; average age: 23.8 years; range: 21 – 29 years) of the University of Kaiserslautern were paid for their participation in the experiment or participated in partial fulfillment of a course requirement. All of them were right-handed and reported normal hearing and none of them had participated in Experiment A. Fourteen of them had no prior musical experience; the others had received an average of 4.4 years of musical training, but had stopped since 6.6 years on average. Therefore, all of them were considered as nonmusicians.

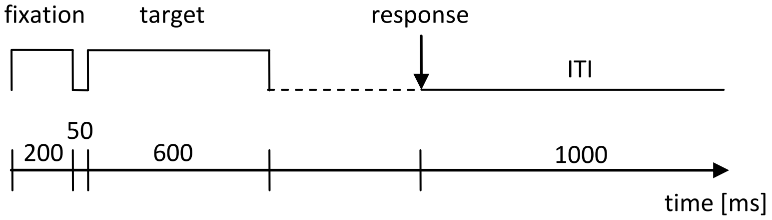

Figure 10: Time course of a trial in the timbre judgment task in Experiment B.

Stimuli and procedure. For this experiment, the stimuli of Experiment A were adopted with the exception that there was no reference tone. The trial-to-trial procedure was the same as in Experiment A except that target sounds followed directly after fixation without a preceding reference tone (see Figure 10). In this experiment, participants had to make judgments based on whether the heard sound was a piano or a vocal tone by pressing a left or right response key. All the other experimental settings remained the same as in Experiment A, resulting in an overall duration of about 20 minutes per participant.

3.2.2.2 Results

Two trials with RT shorter than 100 ms were excluded from analysis. Another 5.1% (197 of 3840 trials, approximately equally distributed across conditions) were removed due to the individual outlier criterion as described in Experiment A. Mean dRT and dER (right – left key) are displayed in Table 3 as a function of timbre, global pitch, and distance from the middle. There was no speed-accuracy tradeoff, as there was no correlation of RT and ER, $r = -.026$, $p = .906$. All analyses were conducted as in Experiment A (for results, see Table 5).

Table 5: Results for the statistical analyses on RT and ER in Experiment B.

	F	p	η_p^2
RT			
Pitch	11.60	.002	.335
Timbre	1.03	.320	.043
Distance	0.08	.787	.003
Pitch × Timbre	5.25	.031	.186
Pitch × Distance	0.14	.707	.006
Timbre × Distance	2.84	.105	.110
Pitch × Timbre × Distance	0.01	.936	.000
ER			
Pitch	21.82	< .001	.487
Timbre	0.75	.394	.032
Distance	1.87	.184	.075
Pitch × Timbre	0.11	.746	.005
Pitch × Distance	0.66	.426	.028
Timbre × Distance	0.10	.757	.004
Pitch × Timbre × Distance	0.88	.359	.037

RT analyses. There was a main effect of Pitch, *F*(1, 23) = 11.60, *p* = .002, displayed by a 5 ms advantage for left hand responses to low tones and a 22 ms advantage for right hand responses to high tones. The interaction of Pitch and Timbre, *F(*1, 23) = 5.25, *p* = .031, revealed the effect of Pitch to only be present for the piano category, *F*(1, 23) = 14.12, *p* = .001, but not for the vocal category, *F*(1, 23) = 1.98, *p* = .173. In the piano category, participants made 2 ms faster left hand responses for low tones and 38 ms faster right hand responses for high tones (see Figure 11). There were no other main effects or interactions. To ensure that none of the observed effects was due to some of the participants having received musical training in their past, another rmANOVA with Former musical training (no training vs. former training) as a between-subjects variable was run in addition. The interaction of Former musical training and Pitch was not significant, *F* < 1, as were all other interactions with this between-subjects factor.

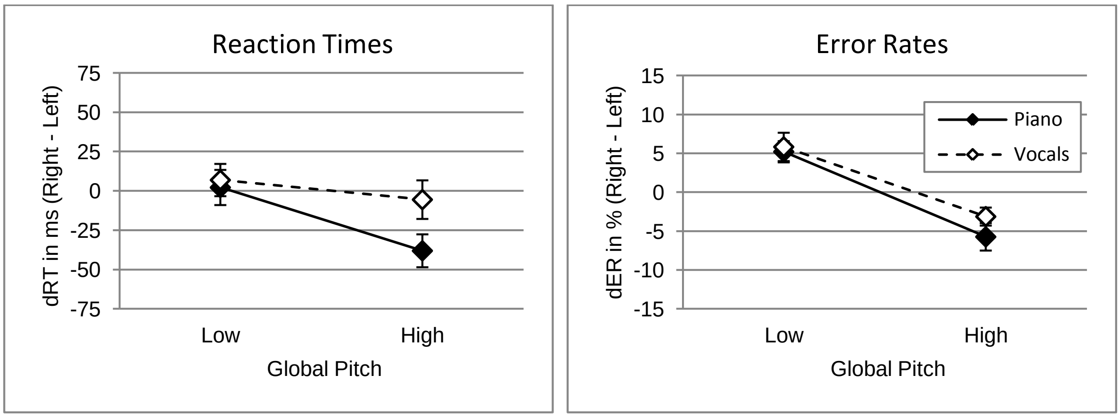

Figure 11: Mean dRT (left panel) and dER (right panel) for the timbre judgment task with small pitch range and horizontally aligned responses in Experiment B.

ER analyses. The rmANOVA on arcsine transformed ER revealed a main effect of Pitch, *F*(1, 23) = 21.82, *p* < .001, confirming the effect found in the RT analysis with a 5.5% advantage for the left/low and a 4.4% advantage for the right/high mapping. None of the other main effects or interactions reached significance.

3.2.2.3 Discussion

The present horizontal timbre judgment task showed that the nonmusician participants made horizontal pitch-to-space associations, regardless of whether they had received any musical training in the past or not. This finding is in agreement with other studies which showed that nonmusicians automatically map pitch onto horizontal space (e.g., Cho et al., 2012; Nishimura & Yokosawa, 2009; Wolf et al., 2012) and contrasts others that did not obtain such effects in similar settings (Lidji et al., 2007; Rusconi et al., 2006).

The major finding in this experiment is that the observed SPARC effect was dependent of timbre. When timbre was the task-relevant feature, participants were only able to ignore the pitch height of the vocal sounds, but not of the piano sounds. This means that, in terms of dimensional overlap, pitch and timbre form a type 7 ensemble, as timbre was not associated with responses. The finding of a pitch-timbre interference in the present experiment goes in line with the observation that - especially for nonmusicians - pitch and timbre interfere in pitch and timbre judgment tasks (Krumhansl & Iverson, 1992; Pitt, 1994). A comparison with the findings of Experiment A, however, shows that the observed interference is asymmetric: In pitch judgments, no such interaction was present while in timbre judgments, interference effects were obtained, which is reflected by the fact that the stimulus sets in both tasks can be allocated within different ensemble types for dimensional overlap (Experiment A: type 5; Experiment B: type 7).

The pitch height of piano sounds was mapped to horizontal space independently of task. This independent mapping implies that for piano tones, pitch is automatically coded spatially. A possible reason for this could be the nature of construction of the piano. On the piano, pitches are ordered ascending from left to right. The knowledge of this organization concept may reinforce the association of low tones with left and high tones with right side responses.

Adopting the horizontal timbre judgment task into a vertical response setting allows for further clarification of the observed findings. Specifically, it provides the opportunity to observe whether vertical spatial mapping of pitch is modulated by timbre as well.

3.2.3 Experiment C: Timbre Judgment in a Vertical Task Setting With Small Pitch Range

Experiment B revealed that, in nonmusicians, automatic pitch-to-space associations were modulated by timbre. Particularly, only when judging the timbre of piano tones, spatial coding of pitch height interfered with response selection.

Experiment C aimed to examine, whether these observations in the horizontal response dimension were caused by similar, timbre-dependent differences in the vertical spatial mapping of pitch heights to responses. Horizontal SPARC effects are attributed to an orthogonal remapping of vertical spatial representations into the horizontal response alignment. Therefore, pitch-timbre interferences in the vertical mapping could modulate horizontal SPARC effects.

To test for such interferences, in Experiment C, task, procedure, and participants of Experiment B were maintained while only the alignment of responses was changed. This enables to conclude whether the findings in Experiment B may have been caused by timbre-related differences in vertically aligned pitch representations.

3.2.3.1 Method

Participants. The 24 students of Experiment B participated in this experiment.

Stimuli and procedure. Stimuli and procedures were adopted from Experiment B with the only change, that response locations were now aligned vertically with a lower and an upper response button (B and 6). Prior studies didn't find any effect of hand assignment to response buttons for nonmusicians (Lidji et al., 2007; Rusconi et al., 2006). In order to not impede the assumed left/low and right/high mapping advantage, a constant assignment of right hand to the upper and left hand to the lower button was maintained.

3.2.3.2 Results

Two trials were excluded from RT analysis (RT < 100 ms). Another 5% (193 of 3840 trials, approximately equally distributed across conditions) were removed due to outlier elimination. Mean dRT and dER (upper – lower key) are displayed in Table 3 as a function of timbre, global pitch, and distance from the middle. There was no correlation of RT and ER, *r* = -.276, *p* = .193, indicating that there was no speed-accuracy tradeoff. Analyses were conducted as in Experiment A, results are displayed in Table 6.

Table 6: Results for the statistical analyses on RT and ER in Experiment C.

	F	p	η_p^2
RT			
Pitch	9.63	.005	.295
Timbre	0.01	.935	.000
Distance	3.19	.087	.122
Pitch × Timbre	0.47	.501	.020
Pitch × Distance	1.26	.273	.052
Timbre × Distance	0.44	.514	.019
Pitch × Timbre × Distance	0.16	.691	.007
ER			
Pitch	7.14	.014	.237
Timbre	2.82	.107	.109
Distance	2.64	.118	.103
Pitch × Timbre	1.01	.325	.042
Pitch × Distance	2.29	.144	.090
Timbre × Distance	4.16	.053	.153
Pitch × Timbre × Distance	0.00	.957	.000

RT analyses. The 2 × 2 × 2 rmANOVA with Pitch, Timbre, and Distance as within-subject variables uncovered a SPARC effect, *F*(1, 23) = 9.63, *p* = .005, which revealed itself through a 10 ms advantage for the low pitch to lower button mapping and a 22 ms advantage for the high pitch to upper button mapping (see Figure 12). None of the other effects reached significance. In order to answer the question whether response dimension affected the timbre dependence of the SPARC effect, another rmANOVA was conducted on the combined data of Experiment B and C with Response alignment (horizontal vs. vertical) as an additional within-subjects factor. Here, as expected, there was a main effect of Pitch, *F*(1, 23) = 17.00, *p* < .001. The only other significant effect was the three-way interaction of Response dimension, Pitch, and Timbre, *F*(1, 23) = 4.86, *p* = .038, confirming that the change of response dimension was the cause for the absence of a Pitch × Timbre interaction in Experiment C.

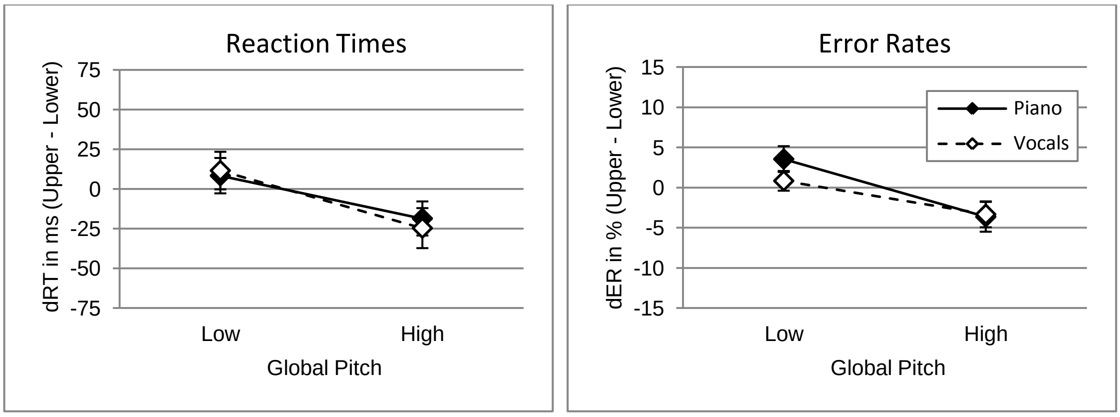

Figure 12: Mean dRT (left panel) and dER (right panel) for the timbre judgment task with small pitch range and vertically aligned responses in Experiment C.

ER analyses. In the error analysis, too, the main effect of Pitch was observed, *F*(1, 23) = 7.14, *p* = .014, with a 2.2% advantage for the compatible low mapping and a 3.5% advantage for the compatible high mapping. None of the other effects reached significance.

3.2.3.3 Discussion

Vertical SPARC was evident regardless of timbre category in the vertical timbre judgment task (see Figure 12). This implies that vertical pitch-to-space associations were formed automatically and independently of timbre. This observation goes in line with those of other studies which reported no effect of timbre on automatic vertical pitch-to-space associations in nonmusicians (e.g., Lidji et al., 2007; Rusconi et al., 2006) and is likely due to the fact that timbre is, if at all, associated with horizontal, but not with vertical space. Therefore, in Experiment C, pitch and timbre form a type 3 ensemble for dimensional overlap.

A comparison between Experiment B and C confirmed that the orientation of the response dimension (i.e., horizontal vs. vertical) is crucial for the observation of pitch-timbre interferences. In particular, pitch-to-space associations for vocal tones are impaired only when responses have to be made along a horizontally aligned response set. This implies that the results in the horizontal timbre judgment task (Experiment B) were not caused by timbre-dependent differences in vertical SPARC.

An alternative explanatory account for the result in Experiment B is introduced by findings in Experiment A. In this horizontal pitch comparison task it was observed that SPARC in vocal sounds interacted with tonal distance from the reference. Specifically, in the vocal category, greater pitch distances from the middle elicited significantly greater dRTs. This means that pitch distances can potentially modulate pitch-to-space associations for vocal sounds.

In order to investigate whether increased pitch differences can affect the automaticity of horizontal mappings of pitch to space, another horizontal timbre judgment task was conducted. In this task, procedures and settings of Experiment B were adopted with a stimulus set with an enlarged pitch range.

3.2.4 Experiment D: Timbre Judgment in a Horizontal Task Setting With Large Pitch Range

The findings in Experiment C revealed that vertical mappings of pitch are generated automatically and independently of timbre. This goes in line with the results of previous studies that did not report timbre differences in their implicit vertical SPARC tasks (e.g., Lidji et al., 2007; Rusconi et al., 2006). The observations made in Experiments B and C allow for the conclusion that the results obtained in the horizontal timbre judgment task (Experiment B) were not caused by pitch-timbre interferences in the vertical setting (Experiment C).

Alternatively, however, pitch representations in implicit horizontal tasks could be modulated by pitch distances. As can be seen in Experiment A, in vocal sounds, tones further away in pitch yielded a significantly larger SPARC effect. Similar observations have also been reported by other studies to the extent that smaller pitch distances caused weaker (Beecham et al., 2009) or even reversed (Rusconi et al., 2006) SPARC effects.

In order to examine possible effects of tonal distances on pitch-to-space associations for vocal sounds, an additional implicit horizontal SPARC task was conducted. In this task, like in Experiment B, participants were required to make timbre judgments with left or right key responses. However, the tone pitches were now separated by larger tonal distances from the middle and each other. This allows exploring whether pitch range can modulate horizontal SPARC in nonmusicians.

3.2.4.1 Method

Participants. Twenty-four students (14 female; average age: 23.5 years; range: 21 – 30 years) of the University of Kaiserslautern were paid for their participation in the experiment or participated in partial fulfillment of a course requirement. All of them were right-handed and reported normal hearing and none of them had participated in any of the other experiments. Fifteen of them had no prior musical experience, the others had received an average of 4.9 years of musical training, but had stopped since 6.3 years on average. Therefore, they were all considered as nonmusicians.

Stimuli and procedure. For this experiment, stimuli with pitch distances of 9 and 12 semitones (F3, G#3, D5, and F5, respectively 174.61, 207.65, 587.33, and 698.46 Hz) from the middle (F4, respectively 349.23 Hz) were produced in two different timbres (vocals and piano). All procedures remained the same as in Experiment B.

3.2.4.2 Results

One of 3840 trials was excluded from analyses because of RTs being smaller than 100 ms. Another 5.3% (204, approximately equally distributed across conditions) of the trials were removed from analyses due to the individual outlier criterion. Table 3 displays mean dRT and dER (right – left key) separately for timbre, global pitch, and distance from the middle. There was no speed-accuracy tradeoff, $r = -.177$, $p = .407$, RT and ER analyses were conducted as in Experiment A (see Table 7 for results).

Table 7: Results for the statistical analyses on RT and ER in Experiment D.

	F	p	η_p^2
RT			
Pitch	38.02	< .001	.623
Timbre	4.59	.043	.166
Distance	0.80	.381	.034
Pitch × Timbre	0.01	.918	.000
Pitch × Distance	3.66	.068	.137
Timbre × Distance	9.53	.005	.293
Pitch × Timbre × Distance	2.06	.165	.082
ER			
Pitch	24.75	< .001	.518
Timbre	1.01	.326	.042
Distance	1.45	.241	.059
Pitch × Timbre	0.01	.922	.000
Pitch × Distance	0.10	.756	.004
Timbre × Distance	4.70	.041	.170
Pitch × Timbre × Distance	.000	.999	.000

RT analyses. The RT analysis revealed a main effect of Pitch, $F(1, 23) = 38.02$, $p < .001$, which expressed itself through a 4 ms advantage for the low pitch/left hand and a 36 ms advantage for the high pitch/right hand response mapping. In addition, a main effect of Timbre was obtained, $F(1, 23) = 4.59$, $p = .043$, with dRTs being on average 27 ms smaller for piano sounds than for vocals (see Figure 13). A careful observation of Table 3 and Figure 13 suggests that, regardless of pitch, piano sounds were generally responded to faster with the right hand. Furthermore, a Timbre × Distance interaction, *F*(1, 23) = 9.53, *p* = .005, was found. Additional analyses revealed the effect of Distance to be significant only for the vocal category, *F*(1, 23) = 8.31, *p* = .008, but not for the piano category, *F*(1, 23) = 0.70, *p* = .410, meaning that in the vocal category, greater distances resulted in average dRTs being 16 ms larger. None of the other effects reached significance. In order to rule out influences of former musical training for the occurrence of the observed main effects and interactions, an additional rmANOVA with the between-subjects factor Former musical training was conducted. Like in Experiment B, the interaction with Pitch, *F* < 1, and also none of the other interactions reached significance.

To verify the influence of global pitch differences on the occurrence of SPARC in the vocal category, an additional 2 × 2 rmANOVA was run on the dRT for vocal sounds with Pitch (low vs. high) as within-subjects variable and Pitch range (6+8 vs. 9+12 semitones) as between-subjects variable in order to compare the results of Experiment B and D. The analysis revealed a main effect of Pitch, *F*(1, 46) = 18.76, *p* < .001, and a two-way interaction of Pitch and Pitch range, *F*(1, 46) = 5.00, *p* = .030, confirming the impact of Pitch range on the occurrence of SPARC in the vocal category.

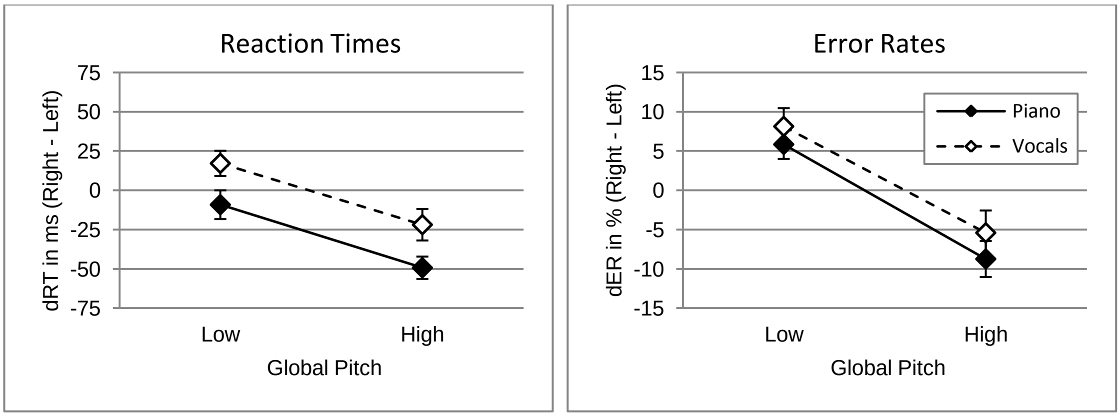

Figure 13: Mean dRT (left panel) and dER (right panel) for the timbre judgment task with large pitch range and horizontally aligned responses in Experiment D.

ER analyses. The main effect of Pitch was also found in the error analyses, *F*(1, 23) = 24.75, *p* < .001, with an advantage of 7.0% for left hand responses to low tones and an advantage of 7.1% for right hand responses to high tones. Again, there was a two-way interaction of Timbre and Distance, *F*(1, 23) = 4.70, *p* = .041. However, the main effect of Distance was significant in the piano category only, *F*(1, 23) = 5.41, *p* = .029, with greater dER for distances closer to the middle.

3.2.4.3 Discussion

In the second horizontal timbre judgment task, horizontal pitch-to-space associations in nonmusicians were observed. In contrast to Experiment B, this SPARC effect was independent of timbre. Furthermore, timbre was associated with horizontal responses. This means that in Experiment D, pitch and timbre form a type 5 ensemble for dimensional overlap, which is different from what was observed in Experiment B (type 7). This difference in results was confirmed by additional analyses which revealed that the SPARC was modulated through increasing the pitch range of the stimulus set. This means that for vocal sounds, when pitch range is increased, automatic mappings of pitch to horizontal responses can be observed.

The observation of an interaction of SPARC and pitch range in the vocal category goes in accordance with results of other studies who reported that SPARC could be modulated by the distance of a sound from the tonal middle of the set of stimuli (Beecham et al., 2009; Rusconi et al., 2006). Specifically, Rusconi et al. (2006) reported that an advantage of the SPARC compatible mapping was obtained only when tones had a distance of at least eight semitone steps from the middle of the tonal range. These results can now be extended with the new finding that pitch distances affect automatic spatial horizontal pitch representations in different ways for different timbres. This, in turn, emphasizes the importance of stimulus set features such as pitch range and timbre categories for the investigation of the horizontal SPARC effect in nonmusicans.

3.3 Joint Discussion

3.3.1 Primary Outcomes

Previous studies on the automatic mapping of pitch to horizontal space in nonmusicians have shown a dependency on task (Cho et al., 2012; Lidji et al., 2007; Rusconi et al., 2006) and the presence of a reference tone (Cho et al., 2012). However, while most of these studies focused on the general presence of SPARC effects in musicians and nonmusicians for different tasks and response settings, little attention has been paid to pitch range and timbre of stimulus sets as possible modulating factors of the horizontal SPARC effect.

Study 1 aimed to investigate the influence of pitch and timbre differences on the automatic spatial coding of pitches. Therefore, four SPARC tasks were conducted in which timbre was controlled for while task relevancy of pitch, orientation of response dimension, and pitch range of the stimulus set were varied. In Experiment A, where pitch height was the task-relevant sound feature, the SPARC effect occurred regardless of timbre. This implies that the participants' spatial coding of pitch height was not modulated by timbre. In this task, timbre and pitch form a type 5 ensemble for dimensional overlap.

In contrast, in Experiment B, where participants had to make judgments based on timbre, the observed automatic mapping of pitch onto horizontal responses was modulated by timbre. In particular, when pitch height was irrelevant for the task, the SPARC effect only occurred for piano tones but not for vocal tones. This implies that the task relevancy of pitch affects interferences between spatial pitch and timbre mappings. In Experiment B, pitch and timbre therefore form a type 7 ensemble.

Experiment C, in which participants had to make timbre judgments with vertically aligned response keys, showed that the findings in Experiment B are specific to the horizontal response dimension. Specifically, when pitch had to be automatically mapped onto vertical space, no influence of timbre category was observed, which reflects the characteristics of a type 3 ensemble. The findings of Experiment B, therefore, cannot be explained by timbre-related differences in vertical pitch representation.

The role of pitch range became evident in Experiment D. In this horizontal timbre judgment task, participants showed a SPARC effect that was not modulated by timbre. Thus, when pitch distances were large enough, also for vocal sounds pitch height was automatically mapped onto horizontal responses. In this experiment, pitch and timbre form a type 5 ensemble.

The results of Study 1 indicate that the pitch range of a stimulus set can affect horizontally aligned spatial mappings of pitch differently, depending on the timbre of a sound, which is reflected by the different types of ensembles that can be ascribed to the same combination of stimulus dimensions: Only in Experiment B, which entailed a timbre judgment task with a small pitch range, the ensemble type reflects pitch and timbre to be integral, inseparable stimulus dimensions, while pitch and timbre are separable in all the other tasks. Specifically, while for piano sounds pitch range does not affect automatic spatial pitch associations along the horizontal dimension, the case is different for vocal sounds, where SPARC was elicited only for tones with a distance of at least 9 semitone steps from the middle of the tonal range.

Altogether, these findings imply that in nonmusicians, stimulus features other than pitch can modulate automatic mapping of pitch height onto horizontal responses. Specifically, pitch processing is not independent of timbre and pitch range in nonmusicians. Note, however, that the interference of pitch and timbre in horizontal response settings is, if present, asymmetrical: While pitch judgments are made independently of timbre, timbre judgments are modulated by pitch. The direction of this interference, therefore, needs to be taken into account when considering the results of bimanual choice reaction tasks involving sounds of different pitch heights and timbres.

3.3.2 Polarity Coding and the Horizontal SPARC Effect

The horizontal SPARC effect is thought to originate from an orthogonal transformation of a vertical pitch-to-space mapping into the horizontal dimension (Cho et al., 2012; Lidji et al., 2007; Nishimura & Yokosawa, 2009; Rusconi et al., 2006). It has been suggested that this orthogonal transformation appears as a result of the advantage for the up-right/down-left mapping of relevant and irrelevant stimulus features to responses in bimanual choice reaction tasks (Cho & Proctor, 2003; Weeks et al., 1995). Specifically, the polarity coding principle can account for both the mapping of pitch to vertical space and for the remapping of this vertical representation into the horizontal plane.

Nonmusicians do not always automatically map pitch heights onto space. Cho et al. (2012) were able to show that nonmusicians, in contrast to musicians, only associated pitch heights with horizontal space automatically, when a reference tone was present. In view of polarity correspondence, the authors reason that in a horizontal response setting, nonmusicians do not automatically code pitch height as polarities along the vertical axis. Cho et al. further argue that only the presence of a reference tone and resulting referential coding of tones as relatively low or high in pitch leads to coding of pitch as polarities that overlap with horizontal responses.

Sounds contain, next to pitch, other stimulus features which can be coded as polarities with the potential to overlap with polarity codes of the response set. This is reflected in the main effect of timbre that was obtained: For both horizontal timbre judgment tasks (i.e., Experiment B and D), an observation of Table 3 suggests a response bias towards right key responses for piano tones. This means that piano sounds are nearly always responded to faster with a right hand response while this was not the case for vocal sounds. The main effect of timbre found in Experiment A and D suggests that the timbre categories were coded as polarities with piano sounds being the + polarity. In Experiment A and D, the effects of timbre and pitch are purely additive and therefore can be considered as independent. In Experiment B, in contrast, timbre interacts with the polarity coding of pitch. Here, in the vocal category, polarity coding of pitch was not sufficient anymore to produce a SPARC effect.

It is likely that in the present experiments, the nature of construction of the piano played a special role. Even if people are not familiar with the ascending left-to-right ordering of pitches on a piano keyboard, they probably know that the keys are horizontally aligned which may in turn facilitate the establishment of polarity codes. Therefore, the results obtained for the horizontal SPARC effect with piano sounds may only reflect pitch-to-space associations for this specific instrument.

The more interesting finding is, therefore, that for sounds of a timbre without such an explicit horizontal association, pitch range modulates horizontal SPARC. Specifically, the findings show that for the vocal sounds used in the present experiments, extending pitch differences and thereby the pitch range modulated the evocation of SPARC. This suggests that when tones are sufficiently distant in pitch from the tonal middle of the stimulus set, pitches can automatically be coded as low and high in a horizontal timbre judgment task. Moreover, a reference tone, compared to which a target is perceived as relatively low or high, is not necessarily required to activate automatic pitch-to-space mappings. This result contrasts the conclusion drawn by Cho et al. (2012) who found a reference tone to be crucial to the evocation of automatic spatial mapping of pitch. However, their results could to some extent be ascribed to their specific choice of artificially produced stimuli which can result in unnatural sounds and impaired timbre identification. For nonmusicians such stimuli could draw more attention to the more salient sound feature timbre and impair the coding of these stimuli according to pitch height. As Cho et al. did not include timbre as a factor in their analysis, the question whether this sound attribute may have affected the horizontal implicit SPARC effect in their sample cannot be answered.

Taken together, the present results indicate that timbre and pitch range can affect the spatial vertical mapping of pitch and the assumed orthogonal transformation of these vertical pitch-to-space mappings into the horizontal dimension. Moreover, the findings imply that both timbre and pitch range should be considered carefully in upcoming research.

4 Study 2: Pitch and Number

4.1 Research Interest

Study 2 aims to investigate whether SNARC and SPARC are independent in the horizontal domain. Therefore, the paradigm of Fischer et al. (2013) was adapted with a set of conditions and tasks more suitable to produce implicit SNARC and SPARC effects. This implies the introduction of variations in stimulus magnitudes (i.e., specific numbers and pitch heights) that are sufficient to produce a SNARC and SPARC effect also in implicit task conditions. Especially for obtaining a more reliable implicit SPARC effect, the results of Study 1 were taken into account for the creation of the stimulus set in Study 2. In addition, a variation of task was introduced. A set of two choice reaction tasks was conducted. In one task (number comparison), numerical magnitude was task-relevant and pitch was not. In the other task (parity judgment), neither numerical magnitude nor pitch height were task-relevant. This means that both SNARC and SPARC were implicit which minimizes the possibility that selective attention to a task-relevant magnitude crowds out the implicit effect.

These experimental tasks offer the opportunity to observe effects of both SNARC and SPARC occurring together. Such a result will be relevant to the issue of the automaticity of these effects. More importantly, however, this will enable to answer the question of their independence. If the effects were to occur as additive factors in the planned factorial design, this would most likely imply that the effects belong to separate mechanisms. Within the taxonomy of the dimensional overlap model (Kornblum et al., 1990; for more detail, see section 2.4.2.2), such an observation would imply that pitch and number form a type 5 ensemble with horizontal responses. An interaction, however, would suggest that spatial numerical and pitch associations belong to joint mechanisms, reflecting a type 8 ensemble. In particular, a size congruity effect of pitch height and numerical magnitude might be observed. If such a size congruity effect was present and specific to tasks involving magnitude judgment, this would allow to study whether observed interferences take place on the decision level of processing, as suggested by the shared decisions account.

4.2 Empirical Section

4.2.1 Method

Participants. Forty-eight students (27 female; average age: 23 years; range: 18 – 27 years) of the University of Kaiserslautern were paid for their participation in the experiment. All of them were right-handed and reported normal hearing. Twenty-eight of them had no formal training in music, the others had received an average of 4.8 years of musical training, but had stopped since 7.3 years on average. Thus they were all considered as nonmusicians.

Stimuli and procedure. Stimuli consisted of the German number words for 1, 2, 8, and 9 (i.e., “eins”, “zwei”, “acht”, and “neun”) that were sung in two low and two high pitches (F3, G#3, D5, and F5) by a proficient female singer, and edited to a duration of 600 ms. The values for the numerical magnitude and pitch height both were chosen according to a 4:3 scheme, which indicates the ratio of the distances around the center. The above mentioned numerical magnitude values are obtained with a center at 5 and distances of 4 and 3; the pitch values with a center at F4 and distances of 12 and 9 semitones. The pitches were specifically chosen as a result of Study 1, which showed that for vocal sounds it is crucial to employ large enough distances in order to obtain a reliable implicit horizontal SPARC effect in nonmusician participants. In addition, clear SNARC effects were reproduced with this specific set of numbers on both number comparison and parity judgment, with numbers sung in one pitch height (F4) only.

Participants were randomly assigned to either of two tasks (magnitude comparison or parity judgment). In the magnitude comparison task, participants heard a single target stimulus and judged whether it was numerically smaller or larger than 5. In the parity judgment task, participants judged if the target number was odd or even. Responses were given along a horizontally aligned response set with a left and right response key (Q and P). In both tasks, participants were instructed to ignore all irrelevant aspects of the numbers.

The experiment was run on Dell laptop computer (Latitude D830, Dell). Stimulus presentation and response collection were controlled through the E-Prime Software (E-Prime 2.0, Psychology Software Tools Inc.). Stimuli were presented via closed Beyerdynamic headphones (DT-770 Pro, Beyerdynamic) at equal loudness well above threshold and clearly within the individual comfort zone.

On each trial, a fixation cross was presented in the centre of the computer screen for a duration of 200ms. 250ms after the onset of the fixation, the target sound was presented for a duration of 600ms. Responses were recorded until 2500ms after stimulus onset, followed by an inter-trial interval of 1000ms.

Participants completed two experimental blocks with 160 trials each (10 repeated measures per stimulus) that only differed in the assignment of categories (magnitude comparison: smaller vs. larger than 5; parity judgment: odd vs. even) to response sides. Trials were presented in a pseudorandom order with the restriction that two consecutive trials did not contain the same target and that there were no more than three consecutive trials with the same correct response location. The order of blocks was counterbalanced across participants. Each block was preceded by a short practice session consisting of 16 trials with feedback after each trial (RT and correctness). Altogether, the experiment lasted about 30 minutes.

4.2.2 Results

RTs shorter than 100 ms were excluded from analysis (6 of 15360 trials). In addition, an individual outlier criterion of two standard deviations was applied to the data of each participant (654 trials, 4.3%, approximately equally distributed across conditions). Mean RT in ms (correct trials only) and ER in % are shown in Table 8 as a function of SNARC compatibility, SPARC compatibility, and task. There was no significant correlation of RT and ER, $p = .70$. A 2 × 2 × 2 rmANOVA with the within-subject variables SNARC compatibility (compatible vs. incompatible) and SPARC compatibility (compatible vs. incompatible) and the between-subjects variable Task (magnitude comparison vs. parity judgment) was conducted on mean RT of correct trials and arcsine-transformed ER (for results, see Table 9).

Table 8: Average RT (ms) and ER (%) as a function of SNARC compatibility, SPARC compatibility, and task.

	magnitude comparison				parity judgment			
	SN c		SN i		SN c		SN i	
	SP c	SP i	SP c	SP i	SP c	SP i	SP c	SP i
RT (ms)	413	424	444	461	459	461	481	485
ER (%)	6.4	7.8	6.4	8.6	6.3	6.4	12.5	15.5

Note. SN c = SNARC compatible; SN i = SNARC incompatible; SP c = SPARC compatible; SP i = SPARC incompatible.

RT analyses. In the RT analysis, the main effect of SNARC compatibility, *F*(1, 46) = 52.54, *p* < .001, was a 25 ms advantage for SNARC compatible trials. The main effect of SPARC compatibility, *F*(1, 46) = 20.04, *p* < .001, was a comparatively smaller advantage of 8 ms for the SPARC compatible trials. The interaction of SPARC compatibility and Task, *F*(1, 46) = 5.92, *p* = .019, revealed the effect of SPARC compatibility only to be present for the magnitude comparison task, *F*(1, 23) = 10.83, *p* < .001, but not for the parity judgment task, *F*(1, 23) = 2.44, *p* = .132 (also see Figure 14).

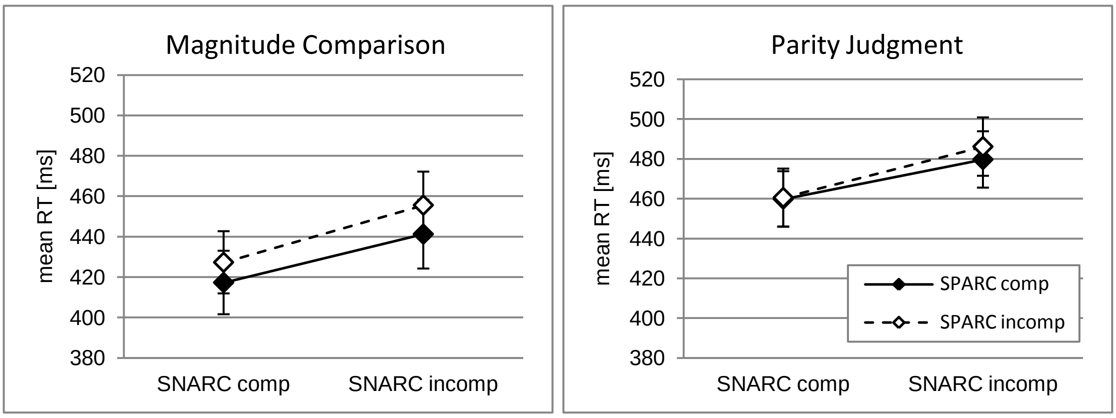

Figure 14: Mean RT for the magnitude comparison (left panel) and the parity judgment (right panel) task as a function of SNARC compatibility, SPARC compatibility, and task.

No other effects reached significance. In particular, the interaction of SNARC and SPARC compatibility, *p* = .204, was far from significance. In principle, an interaction of SNARC and SPARC compatibility could have indicated congruency. In terms of congruency, the SNARC and SPARC compatible trials ('both compatible') and the SNARC and SPARC incompatible trials ('both incompatible') are congruent trials. Congruency effects would therefore lead to a cross over interaction of SNARC and SPARC compatibility. However, no such effect was obtained.

Table 9: Results for the statistical analyses on RT and ER.

	F	p	η_p^2
RT			
SNARC compatibility	52.54	< .001	.533
SNARC compatibility × Task	0.25	.618	.005
SPARC compatibility	20.04	< .001	.303
SPARC compatibility × Task	5.92	.019	.114
SNARC compatibility × SPARC compatibility	1.66	.204	.035
SNARC compatibility × SPARC compatibility × Task	0.03	.859	.001
Task	2.92	.094	.060
ER			
SNARC compatibility	32.11	< .001	.411
SNARC compatibility × Task	23.98	< .001	.343
SPARC compatibility	11.54	< .001	.201
SPARC compatibility × Task	0.04	.836	.001
SNARC compatibility × SPARC compatibility	1.90	.175	.040
SNARC compatibility × SPARC compatibility × Task	1.38	.246	.029
Task	4.34	.043	.086

ER analyses. The rmANOVA on ER was conducted with the same factors as above (also see Figure 15). Again, main effects were obtained of SNARC compatibility, *F*(1, 46) = 32.11, *p* < .001, and SPARC compatibility, *F*(1, 46) = 11.54, *p* < .001. The Task × SNARC compatibility interaction, *F*(1, 46) = 23.98, *p* < .001, indicated that the SNARC effect was present only for the parity judgment task, *F*(1, 23) = 47.41, *p* < .001, but not for the magnitude judgment task, *F*(1, 23) = 0.36, *p* = .550. In addition, there was a main effect of Task, $F(1, 46) = 4.34$, $p = .043$, indicating that parity judgment was less accurate than magnitude comparison.

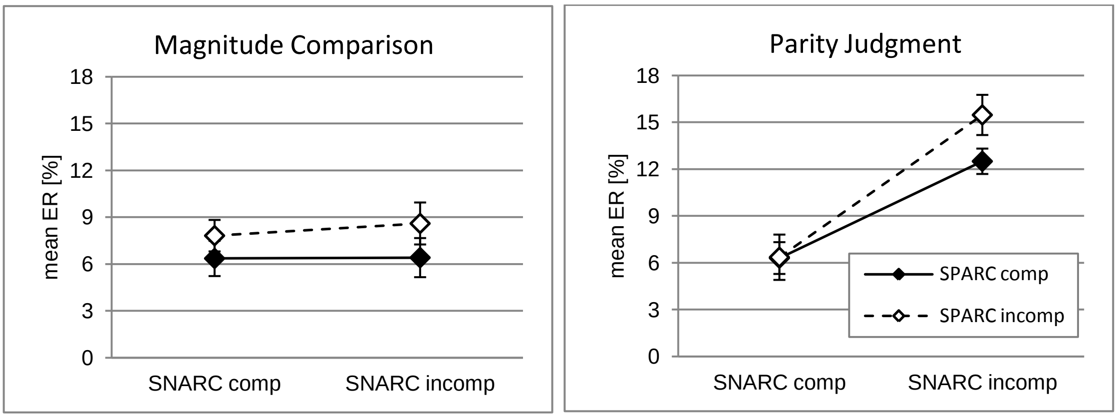

Figure 15: Mean ER for the magnitude comparison (left panel) and the parity judgment (right panel) task as a function of SNARC compatibility, SPARC compatibility, and task.

4.3 Discussion

In order to study whether SNARC and SPARC are independent effects, a factorial design was adapted in which numerical magnitude and pitch height were varied within auditorily presented stimuli (Fischer et al., 2013). These stimuli were employed in a set of two tasks. While in the number comparison task the SNARC effect was explicit and the SPARC effect implicit, in the parity judgment task, both effects were implicit.

In the response times, both SNARC and SPARC effects were found to be present in the magnitude comparison task, even though pitch was not relevant for solving the task. This implies that pitch was automatically coded along with number, or, in other words, that the assessment of numerical magnitude co-activated the spatial coding of pitch. In the parity judgment task, only the SNARC effect was present, while no SPARC effect could be observed. This implies that the spatial code for numbers was activated automatically, while this was not the case for pitch height. The absence of a SPARC effect suggests that pitch height is only coded spatially when magnitude coding is required by the task. However, in the ER, SPARC was present in both tasks. This result points towards the assumption that pitch, like number, is automatically and unconditionally mapped onto space. In sum, it can be concluded that both number and pitch automatically activate a spatial representation.

These results contrast to those of Fischer et al. (2013), who found in both their magnitude judgment tasks that only compatibility of the task-relevant attribute with responses enhanced performance. As mentioned earlier, their results could be attributed to their specific choice of stimuli. For nonmusicians, implicit SPARC is, next to the presence of a reference tone (Cho et al., 2012), sensitive to timbre and pitch range, as shown in Study 1 of the present work. Moreover, an earlier study of Rusconi et al. (2006) showed that for stimuli ranging with tonal differences of 2, 4, and 6 semitone step distances around the middle, vertical SPARC was even reversed in nonmusicians. Therefore, it is likely that part of the stimuli in Fischer and colleagues' set were not sufficient to produce implicit horizontal SPARC effects in nonmusician participants.

Even though both numerical magnitude and pitch height were jointly and automatically mapped to horizontal space, the observed SNARC and SPARC effects were found to be additive and therefore independent. Regarding the dimensional overlap taxonomy (Kornblum et al., 1990), this result means that pitch and number form a type 5 ensemble, where both relevant and irrelevant stimulus dimensions are associated with responses, but not with each other. The additivity of the observed SPARC and SNARC effects includes the absence of congruency effects, and, moreover, their independence implies separate processing mechanisms. In other words, in both tasks employed in the present study, number and pitch appear to be separable dimensions in the sense of Garner's (1974) distinction. A reason for this observation might be that SNARC and SPARC originate on separate representational dimensions (SNARC: horizontal; SPARC: vertical). Another reason could be that, in contrast to other magnitudes for which size congruity effects can be observed (e.g., the joint variation of physical and numerical size within a stimulus set), different verbalizations underlie judgments of the numerical and pitch dimension (*small – large* vs. *low – high*).

Congruency effects would have been expected based on the principle of polarity correspondence (Proctor & Cho, 2006): Potentially, the spatial representations of number and pitch strengthen each other when they share the same polarity and interfere with each other in case their polarities do not match. Because no congruency effects were observed, it is not possible to discuss whether such effects for pitch and number would be task-specific. This would have been expected based on the shared decisions account (Santens & Verguts, 2011), which proposes that the irrelevant dimension will interfere with the decision process, when magnitude judgment is required by the task. Instead, however, the spatial representations of number magnitude and pitch appear to enter the decision stage through separate channels.

5 General Discussion

The focus of the present thesis was to investigate, how associations of pitch with horizontal space are affected by stimulus attributes other than the pitch height of a heard sound, especially when pitch height is not task-relevant. Effects of SRC, like SPARC for pitch heights, can potentially be modulated by different, additionally varied stimulus features, regardless of whether these features are task-relevant or not (e.g., Henik & Tzelgov, 1982; Tlauka, 2002).

Nonmusicians have been shown to produce less stable horizontal pitch-to-space associations (e.g., Cho et al., 2012; Lidji et al., 2007; Rusconi et al., 2006) and generally, findings on horizontal SPARC in nonmusicians are inconsistent. Thus, especially in this participant subgroup, it is of interest whether associations of pitch height with horizontal space may be affected by stimulus features other than pitch.

According to the principle of polarity correspondence (Proctor & Cho, 2006), any kind of structural similarity of stimulus and response set can lead to a facilitation of response selection. This means that ordinal information, but also categorical information, can be coded as polarities that then overlap with responses. Structural similarity is reflected as dimensional overlap in the dimensional overlap model (Kornblum et al., 1990) which allows for more detailed predictions on possible interferences. In order to further clarify the inconsistency of findings on the horizontal SPARC effect in nonmusicians, the present thesis included two separate studies that investigated interactions of horizontal SPARC with stimulus features that contain only categorical (i.e., timbre) or ordinal (i.e., numerical size) information.

5.1 On the Interdependence of SRC Effects for Pitch and Timbre

Earlier studies on horizontal SPARC in nonmusicians mostly focused on the presence of SPARC depending on the task-relevance of pitch and response setting (e.g., Lidji et al., 2007; Rusconi et al., 2006), but little attention was paid to features of the stimulus set. Study 1 aimed to investigate, how pitch and timbre differences within a stimulus set influence the automatic spatial coding of pitches. Therefore, a set of four SPARC tasks was conducted with nonmusicians, in which task relevancy of pitch, orientation of response dimension, and pitch range of the stimulus set were varied while timbre was controlled for.

Results showed that participants' pitch-to-space associations were formed independently of timbre, when pitch was task-relevant in a horizontal response setting. However, the case was different, when pitch was a task-irrelevant stimulus feature. Here, in the horizontal response dimension, pitch-to-space associations were modulated by instrumental timbre. However, this effect of timbre could not be replicated in the vertical response dimension, which leads to the conclusion that the interaction of pitch and timbre in the horizontal timbre judgment task was not caused by timbre-specific differences of associations of pitch and vertical space. A replication of the horizontal timbre judgment task with a widened pitch range yielded SPARC effects independently of timbre and also associations of timbre with response sides. These results showed that for the horizontal SPARC effect with vocal sounds, large enough pitch distances are required for the forming of an automatic mapping of task-irrelevant pitch to horizontal space.

The results of Study 1 indicated that the pitch range of a stimulus set can affect horizontal SPARC, depending on the timbre of a sound. This is particularly the case for vocal sounds, as for this timbre category, SPARC was only present when stimuli spanned over a large tonal range. In sum, pitch processing is not always independent of timbre and pitch range in nonmusicians. Specifically, these factors can affect the assumed orthogonal transformation of vertical pitch-to-space mappings into the horizontal dimension when pitch height is not task-relevant. However, when pitch ranges are chosen sufficiently wide, SRC for pitch and timbre are formed independently. In view of future research, the possible interferences need to be taken into account when conducting choice-reaction tasks involving sounds of varying pitch heights and timbres.

5.2 Joint Processing of Pitch and Number: Horizontal SPARC and SNARC do not Interact

The question of whether SNARC and SPARC are independent has been addressed by earlier studies (Beecham et al., 2009; Fischer et al., 2013). A final answer could not be given, as only an indirect measure was used or the stimulus set was not sufficient to produce isolated implicit SRC effects of pitch and number. In order to provide a more conclusive answer, Study 2 involved a factorial design in which numerical size and pitch height of the stimuli were varied within one and the same stimulus set. The improvement here was that it was ensured that the chosen numbers, as well as the pitch heights, were sufficient to produce isolated explicit and implicit SRC effects in a horizontal task setting. In addition, task relevance of magnitude processing was varied. The new stimuli were employed in a parity judgment and in a number comparison task. Thus, while pitch was always task-irrelevant, numerical magnitude could be task-relevant or not.

Results showed that both pitch height and numerical magnitude were automatically coded spatially in both tasks. However, horizontal SPARC and SNARC did not interact. Specifically, the effects were only additive and can therefore be considered independent. This finding implies that both numerical magnitude and pitch facilitated responses with corresponding responses, and caused, unlike the prediction made by the shared decisions account, no congruency effects that might have been task-specific.

The results of Study 2 show that, with a carefully chosen stimulus set, both horizontal SPARC and SNARC can be observed as implicit effects in tasks that do or do not involve the assessment of a magnitude. The independence of these effects implies that spatial representations of numerical magnitude and pitch height enter the decision stage through separate channels, which might be because the effects originate from separate representational dimensions (i.e., horizontal vs. vertical). In comparison with an earlier study which yielded different results (Fischer et al., 2013), the observations made in Study 2 again emphasize the importance of a careful consideration of stimulus features such as pitch range for similar future studies.

5.3 SRC for Pitch and Simultaneously Perceived Stimulus Features

Study 1 and 2 aimed to investigate whether stimulus features that contain categorical (i.e., timbre) or ordinal (i.e., numerical size) information can interfere with horizontal pitch-to-space associations in nonmusicians. Study 1 showed that pitch processing is not always independent of timbre and pitch range. However, an enlargement of the chosen tonal range within the same timbre categories produced SRC effects for both pitch and timbre, which were independent from each other. Thus, pitch and categorical information (i.e., timbre) did not interact when the pitch range was large enough to produce automatic horizontal pitch-to-space associations in all timbre categories. Study 2 yielded similar results: With pitches chosen from the same large tonal range as in Study 1, horizontal SPARC and SNARC were both observed. Just like in Study 1, SRC for pitch and numerical (or ordinal) information were both present, but independent.

The results obtained in Study 1 and 2 imply that when the pitch range is chosen large enough, horizontal SPARC is independent of other simultaneously perceived stimulus features, even when those additional stimulus features cause SRC effects themselves. This similarity of results is reflected by the type of ensemble that can be ascribed to the stimulus sets used in Study 1 and Study 2 according to the taxonomy of the dimensional overlap model (Kornblum et al., 1990): In Experiment D of Study 1, as well as in Study 2, pitch and the additionally varied stimulus feature formed type 5 ensembles. The results of Nishimura & Yokosawa (2009), who combined horizontal SPARC with an intrinsically spatial feature (i.e., the side of the ear to which a sound was presented), complete the picture: In their study, the authors observed that horizontal SPARC and the Simon effect (Simon & Rudell, 1967) were both present, but did not interact, which means that pitch and presentation side also form a type 5 ensemble for dimensional overlap.

It is important to note that the observed independence of SPARC of other stimulus features seems to be, however, restricted to stimulus features that are originally represented on different spatial axes. This is indicated through a study in which participants had to reply to either the initial pitch or the direction of pitch change within auditorily presented stimuli with top and bottom key presses. Here, SRC effects were present for both initial pitch and pitch change (Walker & Ehrenstein, 2000). More importantly, however, the authors observed enhanced performance, when initial pitch and pitch change corresponded (e.g., when the pitch of a high tone increased). This means that for pitch and pitch change, which are both semantically bound to the vertical response dimension (i.e., tones were initially *low* or *high* and the pitch changed to the tone being *lower* or *higher*), congruency facilitates responses as expected in terms of the size congruity paradigm (Santens & Verguts, 2011). This implies that pitch and pitch change form a type 8 ensemble with vertically aligned responses within the taxonomy of the dimensional overlap model (Kornblum et al., 1990). Whether the observed congruency effect originates from the decision level of stimulus processing, as expected by the shared decisions account (e.g., Santens & Verguts, 2011), cannot be decided based on the results obtained by Walker and Ehrenstein, as both their tasks involved a direct assessment of pitch.

The independence of horizontal SPARC and further stimulus features as observed in the present thesis can be explained by the fact that, while most SRC effects are thought to originate on a horizontal representational dimension, pitch is represented on a vertical axis. The principle of polarity coding (Proctor & Cho, 2006) can be employed to account for these observations: Horizontal SPARC derives from an orthogonal transformation of the vertical pitch-to-space association into the horizontal response dimension. Thus, the originally formed polarity codes correspond with the lower and upper response option and are transformed in order to generate a response. Other stimulus features, like timbre or numerical magnitude, are represented horizontally in the first place, which yields polarity codes corresponding to left and right response options. Therefore, responses formed according to these stimulus features might enter the decision stage through separate channels, which would explain their independence.

Alternatively, the definition of dimensional overlap (Kornblum et al., 1990) can be applied to the finding that horizontal SPARC is independent of simultaneously varied stimulus features. According to Kornblum et al. (1990), dimensional overlap is present when the categories within stimulus and response sets can be mapped onto each other with a homomorphism that maintains operations and relations within a set.

The extent to which dimensional overlap is present depends on the similarity between these relations. In the present thesis, the varied dimensions of stimulus sets were timbre, pitch, and number while the response set was always binary and aligned horizontally. Because pitch is more strongly and consistently associated to the vertical spatial axis while timbre and number are associated to the horizontal axis, the similarity of relations is reduced by default. Furthermore, there is no semantic overlap between categories within the sets. While pitch can be categorized as low or high, number will be judged as small or large and timbre in the present experiments can be judged as piano or vocals. Taken together, these factors may be the reason why no dimensional overlap between pitch and other stimulus features was observed when responses were given along a horizontally aligned response set.

In sum, the findings of the present thesis imply that stimulus features originating from separate representational dimensions are processed independently more likely than features which originate from the same representational dimension, such as for example numerical and physical size or pitch and pitch change. Moreover, the observations made show that when spatial associations for pitch have to be transformed from the original vertical into the horizontal domain, special attention has to be paid to accessory stimulus features in order to not receive data in which effects are occluded due to features that are coded along with or even more saliently than pitch.

6 Summary and Conclusion

The present thesis aimed to investigate, whether associations of pitch with horizontal space in nonmusicians are affected by stimulus features that can produce compatibility effects with horizontally aligned responses themselves. The findings of Study 1 and 2 show that when the tonal range is wide enough, horizontal pitch-to-space associations are formed automatically, independently of whether additional stimulus information is of categorical (i.e., timbre) or ordinal (i.e., number) nature. The observations made in Study 1 also imply, however, that horizontal SPARC in nonmusicians can be affected by timbre when the employed tonal range is small. In view of these findings, future research should carefully consider timbre and pitch differences within sets of auditory stimuli in order to not occlude possible compatibility or congruity effects.

For further studies on the independence of SPARC and SNARC as observed in Study 2, it would be interesting to employ the stimulus set of Study 2 to a task where neither numerical magnitude, nor pitch height, nor number parity are task-relevant. This could for instance be achieved by introducing an additional, task-relevant stimulus feature (e.g., voice timbre: male vs. female; color of a visually presented shape: red vs. green). Furthermore, the collection of psychophysiological measures of participants' reactions to the same stimulus set could shed more light on the question whether the observed independence of pitch and number results of distinct processing areas. This could, for example, be achieved by employing functional magnetic resonance imaging (fMRI). Based on the assumption that a shared magnitude neural code exists which is located in the IPS (Cohen Kadosh et al., 2008c) and the results obtained in Study 2, it would be expected to find higher levels of activity in the IPS when the spatial association of the stimulus is not compatible with the response side, or, in other words, when stimulus and response are coded as different polarities. Furthermore, magnitude comparison, or in this case, number comparison, should yield higher activation in the left IPS (e.g., Nieder & Dehaene, 2009).

In addition, replications of the above studies with trained musicians and also with different timbre categories (e.g., artificial sounds, non-complex sounds such as sinusoids) might be an interesting continuation of this research, as results could give insight into whether implicit SPARC is always affected by timbre when pitch ranges are rather small and whether timbres less similar to (harmonic) instrumental sounds produce the same effects as those employed in the present work.

Thus, while inconsistencies in horizontal SPARC may not be terminally explained by the results obtained in this thesis, the present work extends previous studies by the knowledge that in nonmusicians, horizontal SPARC can be timbre-dependent and that, when the tonal range is chosen appropriately, horizontal SPARC is independent of other stimulus features that cause SRC effects originating from horizontal representations.

7 References

Algom, D., Dekel, A., & Pansky, A. (1996). The perception of number from the separability of the stimulus: The Stroop effect revisited. *Memory & Cognition*, *24*(5), 557–572.

ANSI (1994). *American National Standard Acoustical Terminology*. New York: American National Standards Institute.

Antell, S. E., & Keating, D. P. (1983). Perception of numerical invariance in neonates. *Child Development*, *54*(3), 695–701.

Bächtold, D., Baumüller, M., & Brugger, P. (1998). Stimulus-response compatibility in representational space. *Neuropsychologia*, *36*(8), 731–735.

Barber, P., & O'Leary, M. (1997). The relevance of salience: Towards an activational account of irrelevant stimulus-response compatibility effects. In B. Hommel & W. Prinz (Eds.), *Theoretical issues in stimulus-response compatibility* (pp. 135–172). Amsterdam, The Netherlands: North-Holland.

Beal, A. L. (1985). The skill of recognizing musical structures. *Memory & Cognition*, *13*(5), 405–412.

Beecham, R., Reeve, R. A., Wilson, S. J., & Antonietti, A. (2009). Spatial representations are specific to different domains of knowledge. *PLoS ONE*, *4*(5), e5543.

Ben Nathan, M., Shaki, S., Salti, M., & Algom, D. (2009). Numbers and space: Associations and dissociations. *Psychonomic Bulletin & Review*, *16*(3), 578–582.

Boenke, L. T., Ohl, F. W., Nikolaev, A. R., Lachmann, T., & van Leeuwen, C. (2009). Different time courses of Stroop and Garner effects in perception — An Event-Related Potentials study. *NeuroImage*, *45*(4), 1272–1288.

Bonato, M., Zorzi, M., & Umiltà, C. (2012). When time is space: Evidence for a mental time line. *Neuroscience & Biobehavioral Reviews*, *36*(10), 2257–2273.

Brannon, E. M., Wusthoff, C. J., Gallistel, C. R., & Gibbon, J. (2001). Numerical subtraction in the pigeon: Evidence for a linear subjective number scale. *Psychological Science*, *12*(3), 238–243.

Brugger, P., Knoch, D., Mohr, C., & Gianotti, L. R. R. (2004). Is digit-color synaesthesia strictly unidirectional? Preliminary evidence for an implicitly colored number space in three synaesthetes. *Acta Neuropsychologica*, *2*(3), 252–258.

Caclin, A., Giard, M.-H., Smith, B. K., & McAdams, S. (2007). Interactive processing of timbre dimensions: A Garner interference study. *Brain Research*, 1138, 159–170.

Casasanto, D., Phillips, W., & Boroditsky, L. (2003). Do we think about music in terms of space? Metaphoric representation of musical pitch. In R. Alterman & D. Kirsh (Eds.), *Proceedings of the twenty-fifth annual conference of the cognitive science society* (p. 1323–1323). Mahwah, NJ, USA: Lawrence Erlbaum Assoc Publ.

Case, R., & Okamoto, Y. (1996). The role of central conceptual structures in the development of children's thought. *Monographs of the Society for Research in Child Development*, *61*(1-2), v-265.

Cho, Y. S., Bae, G. Y., & Proctor, R. W. (2012). Referential coding contributes to the horizontal SMARC effect. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(3), 726–734.

Cho, Y. S., & Proctor, R. W. (2003). Stimulus and response representations underlying orthogonal stimulus-response compatibility effects. *Psychonomic Bulletin & Review*, *10*(1), 45–73.

Cohen Kadosh, R., Brodsky, W., Levin, M., & Henik, A. (2008a). Mental representation: What can pitch tell us about the distance effect? *Cortex*, *44*(4), 470–477.

Cohen Kadosh, R., Cohen Kadosh, K., & Henik, A. (2008b). When brightness counts: The neuronal correlate of numerical-luminance interference. *Cerebral Cortex*, *18*(2), 337–343.

Cohen Kadosh, R., & Henik, A. (2006). A common representation for semantic and physical properties. *Experimental Psychology (formerly "Zeitschrift für Experimentelle Psychologie")*, *53*(2), 87–94.

Cohen Kadosh, R., Henik, A., Rubinsten, O., Mohr, H., Dori, H., Van de Ven, V., … Linden, D. E. J. (2005). Are numbers special? The comparison systems of the human brain investigated by fMRI. *Neuropsychologia*, *43*(9), 1238–1248.

Cohen Kadosh, R., Lammertyn, J., & Izard, V. (2008c). Are numbers special? An overview of chronometric, neuroimaging, developmental and comparative studies of magnitude representation. *Progress in Neurobiology*, *84*, 132–147.

Crowder, R. G. (1989). Imagery for musical timbre. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 472–478.

Daar, M., & Pratt, J. (2008). Digits affect actions: The SNARC effect and response selection. *Cortex*, *44*(4), 400–405.

Darwin, C. J. (2005). Pitch and auditory grouping. In C. J. Plack, A. J. Oxenham, & R. R. Fay (Eds.), *Springer handbook of auditory research: v. 24. Pitch. Neural coding and perception* (pp. 278–305). New York: Springer.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*(1-2), 1–42.

Dehaene, S. (2011). *The number sense: How the mind creates mathematics* (Rev. and updated ed). New York: Oxford University Press.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, *122*(3), 371–396.

Deininger, R. L., & Fitts, P. M. (1956). Stimulus-response compatibility, information theory, and perceptual-motor performance. In H. Quastler (Ed.), *Information theory in psychology: Problems and methods* (pp. 316–349). New York, NY, US: Free Press.

Dolscheid, S., Shayan, S., Majid, A., & Casasanto, D. (2011). The thickness of musical pitch: Psychophysical evidence for the Whorfian hypothesis. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 537–542). Austin, TX, US: Cognitive Science Society.

Donders, F. C. (1868/1969). On the speed of mental processes. *Acta Psychologica*, *30*, 412–431.

Dormal, V., Andres, M., & Pesenti, M. (2008). Dissociation of numerosity and duration processing in the left intraparietal sulcus: A transcranial magnetic stimulation study. *Cortex*, *44*(4), 462–469.

Dormal, V., Andres, M., & Pesenti, M. (2012). Contribution of the right intraparietal sulcus to numerosity and length processing: An fMRI-guided TMS study. *Cortex*, *48*(5), 623–629.

Dormal, V., & Pesenti, M. (2007). Numerosity-length interference. *Experimental Psychology (formerly "Zeitschrift für Experimentelle Psychologie")*, *54*(4), 289–297.

Eimer, M., Hommel, B., & Prinz, W. (1995). S-R compatibility and response selection. *Acta Psychologica*, *90*(1-3), 301–313.

Elkin, J., & Leuthold, H. (2011). The representation of pitch in auditory imagery: Evidence from S-R compatibility and distance effects. *Journal of Cognitive Psychology*, *23*(1), 76–91.

Fias, W., & Fischer, M. H. (2005). Spatial representation of number. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 43–54). New York: Psychology Press.

Fias, W., Lammertyn, J., Reynvoet, B., Dupont, P., & Orban, G. A. (2003). Parietal representation of symbolic and nonsymbolic magnitude. *Journal of Cognitive Neuroscience*, *15*(1), 47–56.

Fischer, M. H. (2006). The future for SNARC could be stark. *Cortex*, *42*(8), 1066–1068.

Fischer, M. H., Castel, A. D., Dodd, M. D., & Pratt, J. (2003). Perceiving numbers causes spatial shifts of attention. *Nature Neuroscience*, *6*(6), 555–556.

Fischer, M. H., Mills, R. A., & Shaki, S. (2010). How to cook a SNARC: Number placement in text rapidly changes spatial–numerical associations. *Brain and Cognition*, *72*(3), 333–336.

Fischer, M. H., Riello, M., Giordano, B. L., & Rusconi, E. (2013). Singing numbers… in cognitive space - A dual-task study of the link between pitch, space, and numbers. *Topics in Cognitive Science*, *5*(2), 354–366.

Fischer, M. H., Shaki, S., & Cruise, A. (2009). It takes just one word to quash a SNARC. *Experimental Psychology (formerly "Zeitschrift für Experimentelle Psychologie")*, *56*(5), 361–366.

Fitousi, D., Shaki, S., & Algom, D. (2009). The role of parity, physical size, and magnitude in numerical cognition: The SNARC effect revisited. *Perception & Psychophysics*, *71*(1), 143–155.

Fitts, P. M., & Deininger, R. L. (1954). S-R compatibility: Correspondence among paired elements within stimulus and response codes. *Journal of Experimental Psychology*, *48*(6), 483–492.

Fitts, P. M., & Seeger, C. M. (1953). S-R compatibility: Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology*, *46*(3), 199–210.

Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences*, *4*(2), 59–65.

Garner, W. R. (1974). *The processing of information and structure*. Hillsdale, NJ: Erlbaum.

Gebuis, T., Cohen Kadosh, R., Haan, E., & Henik, A. (2009). Automatic quantity processing in 5-year olds and adults. *Cognitive Processing*, *10*(2), 133–142.

Gevers, W., & Lammertyn, J. (2005). The hunt for SNARC. *Psychology Science*, *47*(1), 10–21.

Gevers, W., Ratinckx, E., de Baene, W., & Fias, W. (2006). Further evidence that the SNARC effect is processed along a dual-route architecture. *Experimental Psychology (formerly "Zeitschrift für Experimentelle Psychologie")*, *53*(1), 58–68.

Gevers, W., Reynvoet, B., & Fias, W. (2003). The mental representation of ordinal sequences is spatially organized. *Cognition*, *87*(3), B87–B95.

Göbel, S. M., Shaki, S., & Fischer, M. H. (2011). The cultural number line: A review of cultural and linguistic influences on the development of number processing. *Journal of Cross-Cultural Psychology*, *42*(4), 543–565.

Hall, J. W., & Peters, R. W. (1981). Pitch for nonsimultaneous successive harmonics in quiet and noise. *The Journal of the Acoustical Society of America*, *69*(2), 509–513.

Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory & Cognition*, *10*(4), 389–395.

Herrera, A., Macizo, P., & Semenza, C. (2008). The role of working memory in the association between number magnitude and space. *Acta Psychologica*, *128*(2), 225–237.

Hoeckner, S., Moeller, K., Zauner, H., Wood, G., Haider, C., Gasner, A., & Nuerk, H.-C. (2008). Impairments of the mental number line for two-digit numbers in neglect. *Cortex*, *44*(4), 429–438.

Holmes, K. J., & Lourenco, S. F. (2012). Orienting numbers in mental space: Horizontal organization trumps vertical. *The Quarterly Journal of Experimental Psychology*, *65*(6), 1044–1051.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*(05), 849–878.

Hommel, B., & Prinz, W. (Eds.). (1997). *Theoretical issues in stimulus-response compatibility*. Amsterdam, The Netherlands: North-Holland.

Ishihara, M., Keller, P., Rossetti, Y., & Prinz, W. (2008a). Horizontal spatial representations of time: Evidence for the STEARC effect. *Cortex*, *44*(4), 454–461.

Ishihara, M., Keller, P., Rossetti, Y., & Prinz, W. (2008b). Spatial representations of time in action. *International Journal of Psychology*, *43*(3-4), 678.

Kahnemann, D., & Henik, A. (1981). Perceptual organization and attention. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 181–211). Hillsdale, NJ: Erlbaum.

Kahnemann, D., & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & R. Davies (Eds.), *Varieties of attention* (pp. 29–61). San Diego, CA, US: Academic Press.

Kemler Nelson, D. G. (1993). Processing integral dimensions: The whole view. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(5), 1105–1113.

Keus, I. M., Jenks, K. M., & Schwarz, W. (2005). Psychophysiological evidence that the SNARC effect has its functional locus in a response selection stage. *Cognitive Brain Research*, *24*(1), 48–56.

Keus, I. M., & Schwarz, W. (2005). Searching for the functional locus of the SNARC effect: Evidence for a response-related origin. *Memory & Cognition*, *33*(4), 681–695.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility--A model and taxonomy. *Psychological Review*, *97*(2), 253–270.

Kornblum, S., & Lee, J.-W. (1995). Stimulus-response compatibility with relevant and irrelevant stimulus dimensions that do and do not overlap with the response. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(4), 855–875.

Krumhansl, C. L., & Iverson, P. (1992). Perceptual interactions between musical pitch and timbre. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(3), 739–751.

Laski, E. V., & Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child Development*, *78*(6), 1723–1743.

Levitin, D. J. (1999). Memory for musical attributes. In P. R. Cook (Ed.), *Music, cognition and computerized sound: An introduction to psychoacoustics* (pp. 209–227). Cambridge, MA: MIT Press.

Lidji, P., Kolinsky, R., Lochy, A., & Morais, J. (2007). Spatial associations for musical stimuli: A piano in the head? *Journal of Experimental Psychology: Human Perception and Performance*, *33*(5), 1189–1207.

Lindemann, O., Abolafia, J. M., Pratt, J., & Bekkering, H. (2008). Coding strategies in number space: Memory requirements influence spatial–numerical associations. *The Quarterly Journal of Experimental Psychology*, *61*(4), 515–524.

Lockhead, G. R. (1964). Effects of dimensional redundancy on visual discrimination. *Journal of Experimental Psychology*, *73*(1), 95–104.

Lockhead, G. R. (1972). Processing dimensional stimuli: A note. *Psychological Review*, *79*(5), 410–419.

Lourenco, S. F., & Longo, M. R. (2009). Multiple spatial representations of number: Evidence for co-existing compressive and linear scales. *Experimental Brain Research*, *193*(1), 151–156.

Lourenco, S. F., & Longo, M. R. (2010). General magnitude representation in human infants. *Psychological Science*, *21*(6), 873–881.

Lu, C.-H., & Proctor, R. W. (1995). The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. *Psychonomic Bulletin & Review*, *2*(2), 174–207.

McAdams, S., Winsberg, S., Donnadieu, S., Soete, G., & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specifities, and latent subject classes. *Psychological Research*, *58*(3), 177–192.

Melara, R. D., & Marks, L. E. (1990a). Interaction among auditory dimensions: Timbre, pitch, and loudness. *Perception & Psychophysics*, *48*(2), 169–178.

Melara, R. D., & Marks, L. E. (1990b). Perceptual primacy of dimensions: Support for a model of dimensional interaction. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(2), 398–414.

Menon, V., Levitin, D. J., Smith, B. K., Lembke, A., Krasnow, B. D., Glazer, D., Glover, … McAdams, S. (2002). Neural correlates of timbre change in harmonic sounds. *NeuroImage*, *17*(4), 1742–1754.

Moore, B. C. J. (2003). *An introduction to the psychology of hearing* (5th ed.). London: Academic Press.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature*, 215, 1519-1520.

Mudd, S. A. (1963). Spatial stereotypes of four dimensions of pure tone. *Journal of Experimental Psychology*, *66*(4), 347–352.

Müller, D., & Schwarz, W. (2007). Exploring the mental number line: Evidence from a dual-task paradigm. *Psychological Research*, *71*(5), 598–613.

Nieder, A., & Dehaene, S. (2009). Representation of number in the brain. *Annual Review of Neuroscience*, *32*(1), 185–208.

Nishimura, A., & Yokosawa, K. (2009). Effects of laterality and pitch height of an auditory accessory stimulus on horizontal response selection: The Simon effect and the SMARC effect. *Psychonomic Bulletin & Review*, *16*(4), 666–670.

Notebaert, W., Gevers, W., Verguts, T., & Fias, W. (2006). Shared spatial representations for numbers and space: The reversal of the SNARC and the Simon effects. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(5), 1197–1207.

Nuerk, H.-C., Iversen, W., & Willmes, K. (2004). Notational modulation of the SNARC and the MARC (linguistic markedness of response codes) effect. *The Quarterly Journal of Experimental Psychology Section A*, *57*(5), 835–863.

Nuerk, H.-C., Wood, G., & Willmes, K. (2005). The universal SNARC effect. *Experimental Psychology (formerly "Zeitschrift für Experimentelle Psychologie")*, *52*(3), 187–194.

Núñez, R. E. (2011). No innate number line in the human brain. *Journal of Cross-Cultural Psychology*, *42*(4), 651–668.

Núñez, R. E., Doan, D., & Nikoulina, A. (2011). Squeezing, striking, and vocalizing: Is number representation fundamentally spatial? *Cognition*, *120*(2), 225–235.

Pansky, A., & Algom, D. (1999). Stroop and Garner effects in comparative judgment of numerals: The role of attention. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(1), 39–58.

Pinel, P., Piazza, M., Le Bihan, D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron*, *41*(6), 983–993.

Pinhas, M., Pothos, E. M., & Tzelgov, J. (2013). Zooming in and out from the mental number line: Evidence for a number range effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(3), 972–976.

Pitt, M. A. (1994). Perception of pitch and timbre by musically trained and untrained listeners. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(5), 976–986.

Plack, C. J., & Oxenham, A. J. (2005). Overview: The present and future of pitch. In C. J. Plack, A. J. Oxenham, & R. R. Fay (Eds.), *Springer handbook of auditory research: v. 24. Pitch. Neural coding and perception* (pp. 1–6). New York: Springer.

Pomerantz, J. R., & Lockhead, G. (Eds.) (1991). *The perception of structure*. Washington, DC: American Psychological Association.

Pratt, C. C. (1930). The spatial character of high and low tones. *Journal of Experimental Psychology*, *13*(3), 278–285.

Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin*, *132*(3), 416–442.

Proctor, R. W., & Reeve, T. G. (1985). Compatibility effects in the assignment of symbolic stimuli to discrete finger responses. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(5), 623–639.

Proctor, R. W., & Reeve, T. G. (Eds.). (1990). *Advances in psychology: Vol. 65*. *Stimulus-response compatibility: An integrated perspective*. Amsterdam, New York, New York, N.Y., U.S.A: North-Holland.

Proctor, R. W., & Vu, K. P. L. (2006). *Stimulus-response compatibility principles: Data, theory, and application.* Boca Raton, FL: CRC Press.

Reeve, T. G., & Proctor, R. W. (1990). The salient-features coding principle for spatial- and symbolic-compatibility effects. In R. W. Proctor & T. G. Reeve (Eds.), *Advances in psychology: Vol. 65. Stimulus-response compatibility. An integrated perspective* (pp. 163–180). New York, N.Y., U.S.A: North-Holland.

Ren, P., Nicholls, M. E. R., Ma, Y.-Y., Chen, L., & Dyer, A. G. (2011). Size matters: Non-numerical magnitude affects the spatial coding of response. *PLoS ONE*, *6*(8), e23553.

Restle, F. (1970). Speed of adding and comparing numbers. *Journal of Experimental Psychology*, *83*(2, Pt.1), 274–278.

Roederer, J. G. (2008). *The physics and psychophysics of music: An introduction* (4th ed). New York: Springer.

Roffler, S. K., & Butler, R. A. (1968). Localization of tonal stimuli in the vertical plane. *The Journal of the Acoustical Society of America*, *43*(6), 1260.

Rosenbloom, P. S. (1986). The chunking of goal hierarchies: A model of stimulus-response compatibility and practice. In J. Laird, P. S. Rosenbloom, & A. Newell (Eds.), *Universal subgoaling and chunking: The automatic generation and learning of goal hierarchies* (pp. 133–282). Boston: Kluwer.

Rosenbloom, P. S., & Newell, A. (1987). An integrated computational model of stimulus-response compatibility and practice. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 1–52). San Diego, CA, US: Academic Press.

Rusconi, E., Kwan, B., Giordano, B. L., Umiltà, C., & Butterworth, B. (2005). The mental space of pitch height. *Annals of the New York Academy of Sciences*, *1060*(1), 195-197.

Rusconi, E., Kwan, B., Giordano, B. L., Umiltà, C., & Butterworth, B. (2006). Spatial representation of pitch height: the SMARC effect. *Cognition*, *99*(2), 113–129.

Russo, F. A., & Thompson, W. F. (2005). An interval size illusion: The influence of timbre on the perceived size of melodic intervals. *Perception & Psychophysics*, *67*(4), 559–568.

Salillas, E., El Yagoubi, R., & Semenza, C. (2008). Sensory and cognitive processes of shifts of spatial attention induced by numbers: An ERP study. *Cortex*, *44*(4), 406–413.

Santens, S., & Gevers, W. (2008). The SNARC effect does not imply a mental number line. *Cognition*, *108*(1), 263–270.

Santens, S., & Verguts, T. (2011). The size congruity effect: Is bigger always more? *Cognition*, *118*(1), 94–110.

Santiago, J., Lupiáñez, J., Pérez, E., & Funes, M. J. (2007). Time (also) flies from left to right. *Psychonomic Bulletin & Review*, *14*(3), 512–516.

Schwartz, D. A., & Purves, D. (2004). Pitch is determined by naturally occurring periodic sounds. *Hearing Research*, *194*(1-2), 31–46.

Schwarz, W., & Heinze, H.-J. (1998). On the interaction of numerical and size information in digit comparison: A behavioral and event-related potential study. *Neuropsychologia*, *36*(11), 1167–1179.

Schwarz, W., & Ischebeck, A. (2003). On the relative speed account of number-size interference in comparative judgments of numerals. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(3), 507–522.

Shaki, S., & Fischer, M. H. (2008). Reading space into numbers – a cross-linguistic comparison of the SNARC effect. *Cognition*, *108*(2), 590–599.

Shaki, S., Fischer, M. H., & Petrusic, W. M. (2009). Reading habits for both words and numbers contribute to the SNARC effect. *Psychonomic Bulletin & Review*, *16*(2), 328–331.

Shaki, S., Petrusic, W. M., & Leth-Steensen, C. (2012). SNARC effects with numerical and non-numerical symbolic comparative judgments: Instructional and cultural dependencies. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(2), 515–530.

Shayan, S., Ozturk, O., & Sicoli, M. A. (2011). The thickness of pitch: Crossmodal metaphors in Farsi, Turkish, and Zapotec. *Senses & Society*, *6*(1), 96–105.

Shepard, R. N. (1964). Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, *1*(1), 54–87.

Shepard, R. N. (1982). Geometrical approximations to the structure of musical pitch. *Psychological Review*, *89*(4), 305–333.

Shoben, E. J., Čech, C. G., Schwanenflugel, P. J., & Sailor, K. M. (1989). Serial position effects in comparative judgments. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(2), 273–286.

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development*, *75*(2), 428–444.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science*, *14*(3), 237–250.

Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In R. W. Proctor & T. G. Reeve (Eds.), *Advances in psychology: Vol. 65. Stimulus-response compatibility. An integrated perspective* (pp. 31–86). New York, N.Y., U.S.A: North-Holland.

Simon, J. R., & Rudell, A. P. (1967). Auditory S-R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology*, *51*(3), 300–304.

Singh, P. G. (1992). Influence of spectral locus and F0 changes on the pitch and timbre of complex tones. *The Journal of the Acoustical Society of America*, *92*(5), 2650.

Starkey, P., & Cooper, R. (1980). Perception of numbers by human infants. *Science*, *210*(4473), 1033–1035.

Starkey, P., Spelke, E. S., & Gelman, R. (1990). Numerical abstraction by human infants. *Cognition*, *36*(2), 97–127.

Stoffer, T. H., & Umiltà, C. (1997). Spatial stimulus coding and the focus of attention in S-R compatibility and the Simon effect. In B. Hommel & W. Prinz (Eds.), *Theoretical issues in stimulus-response compatibility* (pp. 181–208). Amsterdam, The Netherlands: North-Holland.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology: General*, *121*(1), 15–23.

Stumpf, C. (1883). *Tonpsychologie [Psychology of sound]*. Leipzig: Verlag S. Hirzel.

Tagliabue, M., Zorzi, M., & Umiltà, C. (2002). Cross-modal re-mapping influences the Simon effect. *Memory & Cognition*, *30*(1), 18–23.

Trimble, O. C. (1934). Localization of sound in the anterior-posterior and vertical dimensions of "auditory" space. *British Journal of Psychology - General Section*, *24*, 320–334.

Tzelgov, J., Meyer, J., & Henik, A. (1992). Automatic and intentional processing of numerical information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(1), 166–179.

Ueda, K., & Ohgushi, K. (1987). Perceptual components of pitch: Spatial representation using a multidimensional scaling technique. *The Journal of the Acoustical Society of America*, *82*(4), 1193.

van Dijck, J.-P., & Fias, W. (2011). A working memory account for spatial–numerical associations. *Cognition*, *119*(1), 114–119.

van Dijck, J.-P., Gevers, W., & Fias, W. (2009). Numbers are associated with different types of spatial information depending on the task. *Cognition*, *113*(2), 248–253.

van Dijck, J.-P., Gevers, W., Lafosse, C., & Fias, W. (2012). The heterogeneous nature of number-space interactions. *Frontiers in Human Neuroscience*, 5:182.

Vallesi, A., Binns, M. A., & Shallice, T. (2008). An effect of spatial-temporal association of response codes: Understanding the cognitive representations of time. *Cognition*, *107*(2), 501–527.

Vanzella, P., Schellenberg, E. G., & Bishop, D. (2010). Absolute pitch: Effects of timbre on note-naming ability. *PLoS ONE*, *5*(11), e15449.

Verguts, T., & Fias, W. (2008). Symbolic and nonsymbolic pathways of number processing. *Philosophical Psychology*, *21*(4), 539–554.

Verguts, T., & van Opstal, F. (2005). Dissociation of the distance effect and the size effect in one-digit numbers. *Psychonomic Bulletin & Review*, *12*(5), 925–930.

Viarouge, A., Hubbard, E. M., Dehaene, S., & Sackur, J. (2010). Number line compression and the illusory perception of random numbers. *Experimental Psychology (formerly "Zeitschrift für Experimentelle Psychologie")*, *57*(6), 446–454.

Vogel, S. E., Grabner, R. H., Schneider, M., Siegler, R. S., & Ansari, D. (2013). Overlapping and distinct brain regions involved in estimating the spatial position of numerical and non-numerical magnitudes: An fMRI study. *Neuropsychologia*, *51*(5), 979–989.

Vu, K. P. L., Proctor, R. W., & Urcuioli, P. (2003). Transfer effects of incompatible location-relevant mappings on a subsequent visual or auditory Simon task. *Memory & Cognition*, *31*(7), 1146–1152.

Vuilleumier, P., Ortigue, S., & Brugger, P. (2004). The number space and neglect. *Cortex*, *40*(2), 399–410.

Walker, B. N., & Ehrenstein, A. (2000). Pitch and pitch change interact in auditory displays. *Journal of Experimental Psychology: Applied*, *6*(1), 15–30.

Wallace, R. J. (1971). S-R compatibility and the idea of a response code. *Journal of Experimental Psychology*, *88*(3), 354–360.

Warren, J. D., Uppenkamp, S., Patterson, R. D., & Griffiths, T. D. (2003). Separating pitch chroma and pitch height in the human brain. *Proceedings of the National Academy of Sciences*, *100*(17), 10038–10042.

Warrier, C. M., & Zatorre, R. J. (2002). Influence of tonal context and timbral variation on perception of pitch. *Perception and Psychophysics*, *64*(2), 198–207.

Weeks, D. J., Proctor, R. W., & Beyak, B. (1995). Stimulus-Response Compatibility for vertically oriented stimuli and horizontally oriented responses - evidence for spatial coding. *Quarterly Journal of Experimental Psychology Section A - Human Experimental Psychology*, *48*(2), 367–383.

Wolf, A., Bittrich, K., & Blankenberger, S. (2012). Von H bis h'': Der SPARC-Effekt über drei Oktaven [From B2 to B5: The SPARC effect across three octaves]. In A. Bröder, E. Erdfelder, B. E. Hilbig, T. Meiser, R. F. Pohl, & D. Stahlberg (Eds.), *Abstracts of the 54. Tagung experimentell arbeitender Psychologen* (p. 153–153). Lengerich, Germany: Pabst Science Publishers.

Wolpert, R. S. (1990). Recognition of melody, harmonic accompaniment, and instrumentation - musicians vs nonmusicians. *Music Perception*, *8*(1), 95–106.

Wood, G., Willmes, K., Nuerk, H.-C., & Fischer, M. H. (2008). On the cognitive link between space and number: A meta-analysis of the SNARC effect. *Psychology Science*, *50*(4), 489–525.

Zarate, J. M., Ritson, C. R., & Poeppel, D. (2013). The effect of instrumental timbre on interval discrimination. *PLoS ONE*, *8*(9), e75410.

Zhang, H., Zhang, J., & Kornblum, S. (1999). A parallel distributed processing model of stimulus–stimulus and stimulus–response compatibility. *Cognitive Psychology*, *38*(3), 386–432.

Zorzi, M., Priftis, K., Meneghello, F., Marenzi, R., & Umiltà, C. (2006). The spatial representation of numerical and non-numerical sequences: Evidence from neglect. *Neuropsychologia*, *44*(7), 1061–1067.

Zorzi, M., Priftis, K., & Umiltà, C. (2002). Brain damage: Neglect disrupts the mental number line. *Nature*, *417*(6885), 138–139.

Curriculum Vitae

Name: Barbara Estner

2006 - 2011	Studies of psychology with specialization in educational psychology and maths for grammar school teaching at the Catholic University Eichstätt-Ingolstadt (Studium der Psychologie mit schulpsychologischem Schwerpunkt und der Mathematik für das Lehramt an Gymnasien) Thesis (Zulassungsarbeit): „Mathematical achievement and numerosity estimation in grades 3 and 4” (“Mathematikleistung und das Schätzen unbestimmter Mengen in der 3. und 4. Klasse“)
2011	First state examination (1. Staatsexamen)
since 2011	Research scientist at the University of Kaiserslautern, Department of Cognitive and Developmental Psychology

Danksagung

An dieser Stelle möchte ich mich gerne bei einigen Menschen bedanken. Es waren viele Schritte, die mich auf diesem Weg an mein Ziel gebracht haben und ich durfte dabei vielen Menschen begegnen.

Mein großer Dank gilt vor allen anderen meinem Betreuer Thomas Lachmann, der mir die Promotion überhaupt erst ermöglicht hat. Danke für dein Vertrauen in mich, für deine Betreuung und Unterstützung und dein Engagement. Danke, dass du an mich geglaubt hast und Danke für die Freiheit, die ich in dieser Zeit erfahren durfte.

Danken möchte ich auch all meinen Kollegen im Fachgebiet Kognitive und Entwicklungspsychologie, die mich so herzlich und offen in die Abteilung aufgenommen und mir mit vielen kleinen Anregungen und Hilfestellungen immer wieder weitergeholfen haben. Danke an Andreas Schmitt und Steffen Theobald für ihren Einsatz bei verschiedenen Datenerhebungen.

Ganz besonders danken möchte ich schließlich Andrea Prölß, Corinna Christmann und Kirstin Bergström, die mich mit ihrer Unterstützung in allen Lebenslagen durch diese Zeit begleitet haben. Ohne euch wäre es nicht halb so schön gewesen.