MULTIPLE-DIGIT CUE COMBINATION

AN INVESTIGATION OF MULTIPLE-DIGIT CUE COMBINATION:

PSYCHOPHYSICS AND BAYESIAN MODELING

By NINA PRODRIBABA, HON. B.Sc.

A Thesis Submitted to the School of Graduate

Studies in Partial Fulfillment of the Requirements

For the Degree Master of Science

McMaster University
MASTER OF SCIENCE (2018)
Hamilton, Ontario
(Neuroscience)

TITLE: An investigation of optimal multiple-finger cue combination: psychophysics and Bayesian modeling

AUTHOR: Nina Prodribaba, Hon. B.Sc. (McMaster University or Laurier?)

SUPERVISOR: Dr. Daniel Goldreich

NUMBER OF PAGES: x, 29

Abstract

In recent years, computational neuroscientists have suggested that human behaviour, including perception, occurs in a manner consistent with Bayesian inference. According to the Bayesian ideal observer model, the observer combines cues from multiple sensory streams as a weighted average based on each cue's reliability. Most cue-combination research has focused on integration of cues between sensory modalities or within the visual modality. Cue combination within the tactile modality has been relatively rarely studied, and it is still not known whether cues from individual digits combine optimally. In this thesis, we use the ideal observer model to determine whether cues from three different digits are combined optimally. We predicted that cues from multiple digits would be combined according to the optimal cue combination model. To test our hypothesis, we devised a two-interval forced choice (2IFC) task where participants had to discriminate the distal/proximal location of a 1-mm thick edge across the fingerpad(s) of the index (D2), middle (D3), and ring (D4) fingers. We used a Bayesian adaptive method, the ψ method, to compute participants' psychometric functions for single-digit (D2, D3, and D4) and multiple-digit (D23, D24, D34, and D234) conditions. We determined the stimulus level Δx, the distance (mm) between the distal and proximal stimuli locations, at 76% correct probability. This distance corresponds to a sensitivity index $d' = 1$ and is the σ value of the participant's stimulus measurement distribution. We then used the single-digit σ values to predict optimal cue combination for the multiple-digits combinations. We did not observer optimal cue-combination between the digits in this study. We outline potential implications the results of this experimental have on determining how the nervous system combines cues between digits, focusing on theoretical and experimental updates to the experiment that might result in the observation of optimal cue combination between digits.

Preface

This thesis is composed of three chapters. Chapter 1 overviews the background literature in Bayesian inference, tactile perception, and tactile acuity. Chapter 2 is an empirical studying conducted using psychophysics. Chapter 3 is the general discussion where I discuss the findings and implication of the study.

Natural Science and Engineering Research Council of Canada (NSERC) Discovery Grant awarded to Dr. Daniel Goldreich supported the research in this thesis. McMaster University also supported this this through an annual graduate stipend from 2016 to 2018.

Declaration of Academic Achievement

Chapter 2

I was involved in all aspects of empirical research, including the experimental design, programming, data collection, statistical analysis, and writing. My graduate supervisor Dr. Goldreich made major contributions to the experimental design and programming. Farah Hasan, an undergraduate student, assisted in the data collection for the pilot phase of the research (not included in this thesis). Another undergraduate student, Rohoma Zakir, assisted in the data collection for the experiment and an undergraduate thesis student, Enis El-Hamed, assisted in the data collection and statistical analysis of the experiment.

Acknowledgments

First and foremost, I would like to thank my graduate supervisor, Dr. Daniel Goldreich. From our first meeting where we spent hours talking about Bayes Theorem, to all the meetings where we go over the scheduled time, I have been so grateful to have found a mentor that shares my curiosity and passion for human behaviour. Your clever wit, endless patience, and impeccable teaching skills make you the ideal mentor. Thank you for taking me under your wing and showing me that integrals are a form of art, not something to fear. I would like to thank my committee members, Dr. David I Shore and Dr. Jim Lyons, for their time, invaluable advice, and continued support on this incredible journey.

I would also like to thank my former lab mates, Arnav Bharadwaj and Dr. Lux Li, for all the advice you shared with me about your experiences as graduate students. I would also like to thank my current lab mates, Keon Allen, Akash Deep, and Kyle Gauder, for all their helpful advice on my project and the laughter they bring to the lab. My experimental work would not have been possible without the time volunteered by many undergraduate students, to whom I am eternally grateful: Matt, Patrick, Salina, Kailey, and Sabrin.

Table of Contents

CHAPTER 3

List of Figures

Chapter 1

Chapter 2

List of Abbreviations and Symbols

Abbreviations

2IFC	Two-interval force-choice
D2, D3, D4	Digit number 2, 3, and 4
PDF	Probability density function
PSE	Point of subject equality

Symbols

d'	Sensitivity index
ε	Lapse rate
x_i	Stimulus location
Δx	Stimulus level: Distance between distal and proximal stimulus locations
σ	Standard deviation
ψ	Psychometric function

CHAPTER 1

GENERAL INTRODUCTION

How the brain forms a single unified percept using multiple sources of information or cues remains a question frequently investigated by neuroscientists. Much of this research in computational neuroscience has focused on developing theoretical models with the capacity of modelling any perceptual paradigm. One such method used for investigating sensory cue combination is the ideal observer model. In the ideal cue combination model, the observer integrates information from multiple sensory streams via cues as a weighted average of the cues' reliabilities (i.e., inverse variance), minimizing the variance (i.e., maximizing the precision) of the estimate. Theoretically, the optimal cue combination observer model predicts any perceptual paradigm where the cues provide independent information and arise from a common source. In practice, researchers use the ideal observer model most often to estimate the optimal combination of cues between sensory modalities, rarely within a single sensory modality, especially so within the sense of touch, and to our knowledge, not between multiple digits of the hand. Therefore, the purpose of this dissertation is to determine whether participants optimally combine cues from multiple digits. We predicted that participants' performance would be better with multiple digits, as per the optimal cue combination formula. The results of this study will contribute to bridging the gap in knowledge between the sense of touch and other sensory modalities within the realm of optimal cue combination.

The purpose of this chapter is to introduce sensory cue combination by discussing how information is thought to be processed by the brain and how the application of Bayesian inference and probability theory leads to better estimates of participants' performance. Once the framework for cue combination is outlined, the aim of the rest of the chapter will be to demonstrate the lack of tactile studies using the ideal cue combination model and properties of the tactile modality that warrants further investigation into how an observer combines cues from multiple digits.

1.1 Information Processing

How the nervous system processes different types of information it receives from the environment to form a single unified percept remains a question frequently investigated by neuroscientists. This information signals or cues the nervous system about changes in the environment. McCulloch and Pitts (1943) suggested that neural activity is computational. Later, information processing, signal detection theory, was directly applied to perceptual signals, which we now refer to as cues (Green and Swets, 1966). The signal detection theory framework used in psychophysics allows the researcher to compute a participant's threshold, the stimulus level that the participant can detect a certain percentage of the time. The simplest question the researcher can ask the participant is whether they detected a stimulus or not. Signal detection theory recognized that even if you give the same stimulus multiple times, the response will vary, which can be represented as a PDF) represents an observer's internal response to a stimulus (i.e. the observer's internal measurement) as a probability density function (PDF) (Heeger, 1997). The horizontal axis of the PDF is the internal response and the height of the function represents how frequently that level of internal response occurs. Importantly, the width of the PDF indicates that noise or variability is present in the sensory system. That is, the PDF is centered about the true value of the stimulus and typically assumed to be spread normally, where the width of the distribution represents the variability of the response (i.e. the greater the width the less information is carried by cue). For a threshold detection experiment, Signal detection theory assumes that the observer sets a criterion somewhere along the measurement axis that determine whether the internal response is great enough for the observer to respond that the stimulus is indeed present, depending on the utility the observer assigns to both choices (i.e. the stimulus is present or absent). The researcher can manipulate the expected utility the participant assigns to each choice by varying the penalty for incorrect responses and reward for correct responses.

To obtain a performance measure that is independent of an observer's response criterion, researchers often use discrimination rather than detection experiments. In a discrimination experiment, the researcher asks the participant to discriminate between two levels of a stimulus. After repeated trials, we can compute the participant's threshold, which represents the difference between two stimulus levels the participant can discriminate a certain percentage of the time. Researchers often use adaptive methods, such as various staircase methods (Levitt, 1970), to compute the participants' thresholds. However, if a participant reaches the upper or lower limit of the stimulus level, the staircase threshold estimate might not represent the participant's true threshold (Leek 2001). More recently, researchers have adopted more robust adaptive methods to calculate participants' thresholds (Leek 2001). These methods accurately estimate participants' thresholds because they are resistant to outliers. That is, even if a participant reaches the upper or lower limits of the stimulus level, a problem the staircases methods cannot solve, these algorithms can still accurately estimate a participant's threshold.

The information the nervous system receives cues the observer about changes in the environment. With the aid of signal detection theory, we can measure a participant's performance without having to know the participant's internal criterion. Given that the nervous system processes cues from the environment according to the theoretical framework of signal detection theory, how does the observer use this information to form a unified percept?

1.2 Inference and the Inverse Problem

Like the researcher infers a participant's psychometric function based on the data collected during the experiment, the observer also use inference to form a percept. Estimating properties of the environment usually requires observation of multiple sources of information or cues. Cues allow the observer to perceive the size of objects and their distance. When the nervous system detects a cue, the cue provides the observer with information regarding an environmental property. While cues provide relatively perfect information about changes in the environmental, the nervous system has to process the cue, which results in an imperfect estimate by the time the original information carried by the cue. The first step in the process is when a sensory system senses a cue. The conversion of the physical stimulus into electrical potential adds noise to the original information (Beaudry & Renner, 2011). The addition of noise to the cue occurs at many levels of processing. How then is the nervous system able to form accurate percepts? Much empirical research suggests that the nervous system uses probabilistic inference in various ways to reduce its uncertainty about the information it receives (for a review see Fetsch et al., 2013).

For example, in Book VII of his *Republic*, Plato constructs a scene where prisoners are being held captive in an underground den in such a way that they cannot look behind them but must use the shadows on the wall in front of them to reconstruct what is happening behind them. Naturally, it is quite difficult to know what is causing a shadow as there exists multiple scenarios that could give rise to a particular shadow. Even if the prisoner sees a shadow that seems clearly to be that of a person, it would be extremely difficult to know which shadow belongs to which person without some additional information. The prisoner might, for example, tell if the shadow has long hair or is wearing pants, but he would still be uncertain about which shadow belongs to which person. The 3D world collapses across one dimension and presents visual cues in 2D on the retina, obscuring the orientation of the source object and rendering perception uncertain when observing the shadow. Regardless of the sensory system involved, cues are generally open to multiple interpretations. Consequently, the problem of inferring a cause from its observable effects, also known as the inverse problem, is one that the brain must constantly solve.

One solution to the inverse problem is Bayes' Theorem (see Appendix A for derivation). First derived by Thomas Bayes in 1763 in his *An Essay Towards Solving a Problem in the Doctrine of Chance*, Bayes' Theorem plays a pivotal role in modeling participants' perceptual performance and decision making. Bayes' Theorem,

$$P(H_i|D) = \frac{P(D|H_i) \cdot P(H_i)}{\sum_{k=1}^{N} P(D|H_k) \cdot P(H_k)}, \tag{1.1}$$

calculates the posterior probability $P(H_i|D)$ of a particular hypothesis H_i as the product of its prior probability $P(H_i)$ and likelihood $P(D|H_i)$ in light of some data D, normalized by the sum of all products. Many computational neuroscientists believe that both conscious and nonconscious inference occurs in a manner consistent with the Bayesian Observer Model (Figure 1).

In the Bayesian observer model, the stimulus triggers a cascade of events, culminating in an action that can be used as data for estimating a participant's performance. There is some evidence to suggest that perception is mediated by the observer and the decision about the percept is computed by the decision maker (Trommershauser et al., 2011). The main difference might be the context. That is, if participants are required to categorize novel objects, adding a familiar feature to the novel object improves performance (Sims & Colunga, 2011). How the observer and decision-maker interact to influence the action is still not know and is beyond the scope of this thesis. However, we can now discuss potential models that can predict how the brain integrates multiple cues.

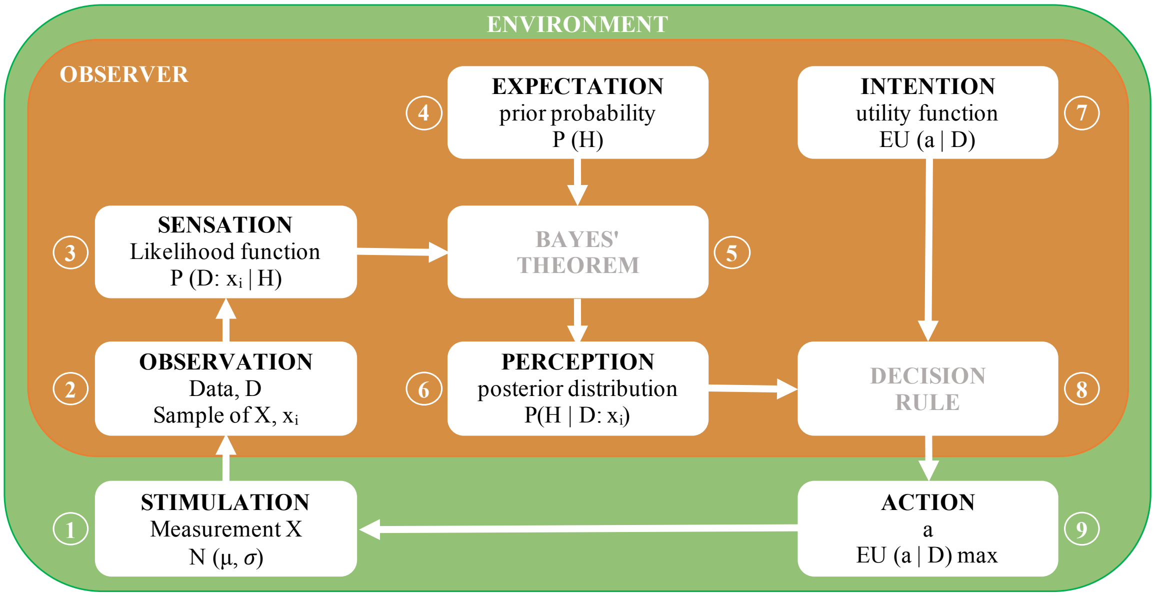

Figure 1. The Bayesian Observer Model. A stimulus (1) results in a sensory measurement x_i sampled from a measurement probability density function centered about the stimulus with Gaussian noise (2), allowing the observer to generate a sensation as the likelihood of the observed data over hypothesized values of the stimulus measurement (3). The observer has some expectation about the hypothesized values of the stimulus measurement based on previous experience with similar stimuli (4). The observer then uses Bayes' Theorem (5) to generate a posterior distribution for the probability each hypothesis given the data it observed (6). The observer uses the posterior distribution and a utility function (7) to generate a decision (8), choosing the action with the maximum expected utility (9). Once the observer executes the action, the observer may get additional feedback from the environment.

1.3 Cue Combination

Given that Bayes' Theorem allows us to solve the inverse problem, we can start to solve more complex perceptual phenomena, such as how the nervous system integrates cues from multiple sensory streams. Continuing with Plato's prisoner example, imagine a single person standing right behind a single prisoner with nobody else around. The prisoner can hear the person talking and he is fairly certain that the sound is coming from directly behind him. To his surprise, though, the shadow he sees on the wall appears to be more to the left of him compared to where he thinks the sound is coming from. Given the conflicting observations, where should the prisoner perceive the person's location? Initial attempts to explain how the nervous system determines what sensation caused a particular percept would suggested that the prisoner rely on the visual system, according to the modality appropriateness hypothesis (Welch & Warren, 1980), which suggested that vision has a greater influence on spatial judgments than hearing, and touch and hearing have a greater influence on temporal judgments compared to vision. Decades later, Alais and Burr (2004a), concluded that the formation of a multisensory percept is determined by the uncertainty of each individual modality, suggesting that perhaps the prisoner should sway more towards the auditory location than the more uncertain visual representation on the wall. More recent multisensory studies have used linear models, particularly the optimal cue combination model, to predict multisensory cue integration.

The general concept of cue combination is that when the brain has multiple sensory inputs to process, it will combine the independent sensory inputs as a weighted average of the stimulus measured values. The resulting combined percept varies depending on the reliability of each cue, approaching the most reliable sense proportionally wi difference in reliability between senses. According to the optimal cue combination formula,

$$\hat{x} = \sum_{i=1}^{n} w_i x_i, \tag{1.2}$$

the perceived stimulus $\hat{x}$ of n independent, Gaussian random variable X_i samples $x_i, i = 1, \dots, n$ that have the same mean μ and variance σ_i^2 is a weighted average, where the weight w_i of cue i is proportional to that cue's reliability r_i, defined as its inverse variance ($r_i = 1/\sigma_i^2$),

$$w_i = \frac{r_i}{\sum_{j=1}^{n} r_j} \tag{1.3}$$

Most of optimal multisensory cue combination research has focused on the integration of visual cues with auditory cues (Battaglia et al., 2003; Alais & Burr, 2004a; Alais & Burr, 2004b; Burr, Banks & Morrone, 2009; Hartcher-O'Brien et al., 2014) and with haptic cues (Ernst & Banks, 2002; Gepshtein & Banks, 2003; Rosas et al., 2005; Takahashi, Diedrichsen & Watt, 2009; Drewing et al., 2008; Kuschel et al., 2010). While limited within-modal optimal cue combination research exists, most of it focuses on depth cues within the visual modality (Jacobs, 1999; Landy & Kojima, 2001; Knill & Saunders, 2003; Hillis et al., 2004; Wismeijer et al., 2010). The handful of optimal cue combination studies conducted within the sense of touch have looked at cues that contribute to the perception of compliance (Bergmann, Tiest & Koppers, 2009) and length perception (Debats et al., 2009). However, the optimal cue combination model, to our knowledge, has not been applied to location discrimination (spatial acuity) within the tactile modality.

1.4 The Sense of Touch

While it is true that a stimulus is better perceived by the observer when the stimulus properties are transmitted through multiple sensory modalities, how is information carried by each sensory modality? In the visual system, retinal photoreceptors, the cones and rods, convert photons into electrical potentials that the nervous system uses to construct a visual image (Kolb, 2003). In the auditory system, tiny hair cells in the cochlea convert sound vibrations into electrical signals (Hudspeth, 1989). In the tactile modality, mechanoreceptors convert mechanical properties of stimuli into electrical potentials the nervous system can use to determine where the stimulus is on the skin (Johnson, 2001). Besides the fact that the sense of touch is the first to develop (Gallese & Ebisch, 2013) and that the tactile sensory organ, the skin, is the largest sensory organ in the human body (Plude, 1987), the receptors used to detect tactile stimuli also innervate other sensory organs. The interconnectivity between receptors of the somatosensory system is complex and remains a topic of research for many neuroscientists. Despite that our sense of touch allows us to have direct contact with our environment, compared to the vision, touch is relatively understudied.

A great deal of work conducted on the sense of touch has focused on spatial acuity, specifically, spatial acuity in the digits through tasks involving passive touch. Craig and colleagues have contributed to determining factors that affect spatial acuity (Craig & Kisner, 1998; Craig 1999) and patterns of temporal order judgment (Carig & Qian, 1997; Craig, 2003). Haptic perception, involving active touch has focused on surface texture, specifically roughness (Lederman & Klatzky, 2009). However, how the digits work together to perform different tasks is still not well understood, making the digits an ideal place to study optimal cue combination within the sense of touch.

There is some evidence to suggest that the information carried by each digit are processes independently. Braido and Zhang (2004) recorded movement of reflexive markers placed at precise joints of the fingers during cylinder grasping and voluntary flexion. Their results showed stereotypical patterns, including proximal-to-distal flexion sequence and characteristic independence. The former results indicate that the initial angle and final angle on the hyperbolic tangent function account for 98% of the variance in the cylinder grasping task, suggesting a high level of synergy from initiation of the grasp to making contact with the cylinder. The latter results indicated an involuntary joint flexion that is induced by a voluntarily flexed joint, suggesting that the isolated movement of one digit will interdependently affect the movement of the adjacent digit. While the movement of the digits appears to depend on the movement of adjacent digits, afferents adjacent to a passively moved finger do not facilitate movement detection (Refshauge, Collins & Gandevia, 2003). Together, these studies suggest that while the digits work together during active touch, the information the nervous system receives from each digit is independent, making the digits the ideal candidates for testing tactile within-modal optimal cue combination.

The complexity of the somatosensory system, which itself is sometimes thought to be multisensory, and the lack of research on the sense of touch compared to other sensory modalities like vision, provides justification for studying optimal cue combination within the sense of touch. Therefore, the purpose of this thesis was to determine whether cues from multiple digits are combined optimally. We hypothesized that cue integration between the digits would be predicted by the optimal cue combination formulae. To test our hypothesis, we devised a two-interval forced choice (2IFC) task where participants had to discriminate the distal/proximal location of 1-mm thick edges across the fingerpads of the index (D2), middle (D3), and ring (D4) fingers of the right hand. We predicted that participants' performance would get better with more digits.

1.5 References

Alais D, Burr D (2004a) The ventriloquist effect results from near-optimal bimodal integration. Curr Biol 14:257–262.
Alais D, Burr D (2004b) No direction-specific bimodal facilitation for audiovisual motion detection. Brain Res Cogn Brain Res 19:185–194.
Battaglia PW, Jacobs RA, Aslin RN (2003) Bayesian integration of visual and auditory signals for spatial localization. JOSA A 20:1391–1397.
Bayes T, Price R, Canton J (1763) An essay towards solving a problem in the doctrine of chances.
Beaudry NJ, Renner R (2011) An intuitive proof of the data processing inequality.
Bergmann Tiest WM, Kappers A (2009) Cues for Haptic Perception of Compliance. IEEE Trans Haptics 2:189–199.
Braido P, Zhang X (2004) Quantitative analysis of finger motion coordination in hand manipulative and gestic acts. Hum Mov Sci 22:661–678.
Burr D, Banks MS, Morrone MC (2009) Auditory dominance over vision in the perception of interval duration. Exp Brain Res 198:49–57.
C. Craig JMK, James (1998) Factors affecting tactile spatial acuity. Somatosens Mot Res 15:29–45.
Craig JC (1999) Grating orientation as a measure of tactile spatial acuity. Somatosens Mot Res 16:197–206.
Craig JC (2003) The effect of hand position and pattern motion on temporal order judgments. Perception & Psychophysics 65:779–788.
Craig JC, Qian X (1997) Tactile pattern perception by two fingers: temporal interference and response competition. Percept psychophys 59:252–265.
Debats NB, van de Langenberg RW, Kingma I, Smeets JB, Beek PJ (2010) Exploratory movements determine cue weighting in haptic length perception of handheld rods. J Neurophysiol 104:2821–2830.
Drewing K, Wiecki TV, Ernst MO (2008) Material properties determine how force and position signals combine in haptic shape perception. Acta Psychol (Amst) 128:264–273.
Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. Nature 415:429.
Fetsch CR, DeAngelis GC, Angelaki DE (2013) Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. 14:429.
Gallese V, Ebisch S (2013) Embodied simulation and touch: The sense of touch in social cognition. Phenomenology and Mind:196–210.
Gepshtein S, Banks MS (2003) Viewing Geometry Determines How Vision and Haptics Combine in Size Perception. Curr Biol 13:483–488.
Hartcher-O'Brien J, Di Luca M, Ernst MO (2014) The duration of uncertain times: audiovisual information about intervals is integrated in a statistically optimal fashion. PLoS One 9:e89339.
Heeger D (1998) Signal detection theory. In: California.

Hillis JM, Watt SJ, Landy MS, Banks MS (2004) Slant from texture and disparity cues: optimal cue combination. J Vis 4:967–992.
Hudspeth AJ (1989) How the ear's works work. Nature 341:397.
Jacobs RA (1999) Optimal integration of texture and motion cues to depth. Vision research 39:3621–3629.
Johnson KO (2001) The roles and functions of cutaneous mechanoreceptors. Curr Opin Neurobiol 11:455–461.
Knill DC, Saunders JA (2003) Do humans optimally integrate stereo and texture information for judgments of surface slant? Vision Research 43:2539–2558.
Kolb H (2003) How the retina works: Much of the construction of an image takes place in the retina itself through the use of specialized neural circuits. American scientist 91:28–35.
Kuschel M, Di Luca M, Buss M, Klatzky RL (2010) Combination and Integration in the Perception of Visual-Haptic Compliance Information. IEEE Trans Haptics 3:234–244.
Landy MS, Kojima H (2001) Ideal cue combination for localizing texture-defined edges. JOSA A 18:2307–2320.
Lederman SJ, Klatzky RL (2009) Haptic perception: a tutorial. Atten Percept Psychophys 71:1439–1459.
Leek MRJP, psychophysics (2001) Adaptive procedures in psychophysical research. 63:1279–1292.
Levitt H (1971) Transformed up-down methods in psychoacoustics. The Journal of the Acoustical society of America 49:467–477.
McCulloch WS, Pitts WJ Tbomb (1943) A logical calculus of the ideas immanent in nervous activity. 5:115–133.
Plude DJ (1987) Sensory, perceptual, and motor function in human aging. The Elderly as modern pioneers:94.
Read JC (2012) Visual perception: understanding visual cues to depth. Curr Biol 22:R163–165.
Refshauge KM, Collins DF, Gandevia SC (2003) The detection of human finger movement is not facilitated by input from receptors in adjacent digits. The Journal of Physiology 551:371–377.
Rosas P, Wagemans J, Ernst MO, Wichmann FA (2005) Texture and haptic cues in slant discrimination: reliability-based cue weighting without statistically optimal cue combination. JOSA A 22:801–809.
Shorey P (1935) The Republic: W. Harvard University Press.
Sims C, Colunga E (2011) Perceptual and Conceptual Cues in Classification and Inference Tasks. In: Proceedings of the Annual Meeting of the Cognitive Science Society.
Takahashi C, Diedrichsen J, Watt SJ (2009) Integration of vision and haptics during tool use. J Vis 9:3 1–13.
Trommershauser J, Kording K, Landy MS (2011) Sensory cue integration: Oxford University Press.

Welch RB, Warren DH (1980) Immediate perceptual response to intersensory discrepancy. Psychological bulletin 88:638.
Wismeijer DA, Erkelens CJ, van Ee R, Wexler M (2010) Depth cue combination in spontaneous eye movements. J Vis 10:25.

CHAPTER 2

MULTIPLE-DIGIT CUE COMBINATION

2.1 Abstract

Many computational neuroscientists believe that human perception occurs in a manner consistent with probabilistic principles. Evidence to support this belief suggests that the nervous system performs like an ideal observer, that is, it integrates cues by calculating a weighted average, weighting each cue by its reliability (i.e. inverse variance). Most cue-combination research has focused on integration of cues between sensory modalities or within the visual modality. Cue combination within the tactile modality has been relatively rarely studied, and it is still not known whether cues from individual digits combine optimally. To answer this question, we devised a two-interval forced choice (2IFC) task where participants had to discriminate the proximal/dorsal location of 1 mm-thick edges across the fingerpads of their index (D2), middle (D3), and/or ring (D4) fingers. We used a Bayesian adaptive algorithm to adjust the stimulus level (the difference between the distal and proximal location in mm) from trial-to-trial based on the minimization of expected entropy of the posterior distribution over psychometric function. Using this algorithm, we obtained each participant's 76-%-correct threshold (i.e. the standard deviation of the participant's internal spatial measurement distribution). We compared the 76%-correct thresholds in the combined-digit conditions with those from the individual digit conditions. We did not observe optimal cue-combination between the digits in this study. We outline potential implications the results of this experimental have on determining how the nervous system combines cues between digits optimally, focusing on theoretical and experimental updates to the experiment.

2.2 Introduction

Most computational neuroscientists believe that the nervous system performs in a similar manner to that of an ideal observer model. The ideal observer combines cues from multiple sensory streams as a weighted average based on each cue's reliability. Most of optimal multisensory cue combination research has focused on the integration of visual cues with auditory cues (Battaglia et al., 2003; Alais & Burr, 2004a; Alais & Burr, 2004b; Burr, Banks & Morrone, 2009; Hartcher-O'Brien et al., 2014) and with haptic cues (Ernst & Banks, 2002; Gepshtein & Banks, 2003; Takahashi, Diedrichsen & Watt, 2009; Drewing et al., 2008; Kuschel et al., 2010). While limited within-modal optimal cue combination research exists, most of it focuses on depth cues within the visual modality (Jacobs, 1999; Landy & Kojima, 2001; Knill & Saunders, 2003; Hillis et al., 2004; Wismeijer et al., 2010; Read, 2012). The handful of optimal cue combination studies conducted within the sense of touch have looked at cues that contribute to the perception of compliance (Bergmann Tiest & Koppers, 2009) and length perception (Debats et al., 2009). However, the optimal cue combination model, to our knowledge, has not been applied to location discrimination (spatial acuity) within the tactile modality.

A great deal of research within the sense of touch has focused on determining how participants perform when they use a single digit, specifically, factors that affect spatial acuity (Craig & Kisner, 1998; Craig 1999) and patterns of temporal order judgment (Carig & Qian, 1997; Craig, 2003). In our lab, we have focused on the grating orientation task (GOT). Among other findings, the lab's GOT studies have revealed that passive tactile spatial acuity on the index finger depends on the size of the digit, possibly because mechanoreceptor density is related inversely to finger surface area (Peters et al., 2009; Wong et al., 2013). In comparison, haptic studies have focused on perception of surface texture (Lederman & Klatzky, 2009). However, the optimal cue combination model has not been applied to either passive or active touch.

Using the ideal observer model as the theoretical framework, the aim of the current study was to bridge the gap between the application of the optimal cue combination model to cues across multiple modalities and within a single modality, as well as to bridge the gap between active and passive touch. To determine whether cues from multiple digits are combined optimally, we devised a two-interval forced choice (2IFC) task in which participants had to discriminate the distal/proximal location of a 1-mm thick edge across the fingerpad(s) of the index (D2), middle (D3), and ring (D4) fingers. We used a Bayesian adaptive method to estimate participants' psychometric functions for single-digit (D2, D3, and D4) and multiple-digit (D23, D24, D34, and D234) conditions. We determined the stimulus level Δx, the distance (mm) between the distal and proximal stimuli locations, at 76% correct probability. This distance corresponds to a sensitivity index $d' = 1$ and is the σ value of the participant's stimulus measurement distribution. We then used the single-digit σ values to predict optimal cue combination for the multiple-digit combinations.

2.3 Materials and Methods

2.3.1 Participants. Twenty-four participants (4 males and 20 females) age 18 to 24 years took part in this study. Five of the participants were left-hand dominant and one ambidextrous. We recruited participants from McMaster University using the PNB Research Participation System (SONA) with the following exclusion criteria: diabetes, nervous system disorder or injury (tremor, epilepsy, multiple sclerosis, stroke, etc.), learning disability, dyslexia, attention deficit disorder, cognitive impairment, carpal tunnel syndrome, arthritis of the hands, hyperhidrosis. We screened against these conditions because they can adversely affect somatosensory neurons (Hyllienmark et al., 1995) or tactile performance (Grant et al., 1999). All participants gave signed consent and received monetary compensation or course credit for their participation. The McMaster University Research Ethics Board approved all procedures used in this experiment.

2.3.2 Apparatus. The stimuli consisted of seven 1-mm thick edges for all combinations of the index (D2), middle (D3), and ring (D4) fingers, including single (D2, D3, and D4)- and multiple (D23, D24, D34, D234)-digit combinations, that we 3D-printed on a discriminandum wheel (Figure 2A). For all 3D-printed components of the apparatus, we used OpenSCAD (version 2015-03-3) to design the components, Cura (version 2.5.1) to prepare the components for printing, and an Ultimaker 3D-printer (product Ultimaker 2 Go) to print the components (PLA plastic). The diameter of the discriminandum wheel was 113 mm, and its height was 80 mm. All edges were 20 mm wide (and 1-mm thick) with different heights: 20 mm for D2, D3, and D4; 50 mm for D23 and D34; two 20-mm edges separated by 40 mm for D24; and 80 mm for D234. We attached the discriminandum wheel to a 73-mm long 3D-printed hexagonal prism, the post, (Figure 2B) of 19.85-mm diameter, which went through a hole in the table. One end of the post fit into a 20-mm diameter hole on the bottom of the discriminandum wheel. The other end was a 3-mm thick disk of 32-mm diameter with four 3-mm diameter holes. We used 4 bolts and nuts to connect the disk end of the post to a custom-milled metal cylinder. We then attached the metal cylinder to the shaft of a stepper motor (NEMA 23, 3-stack, Industrial Devices). Finally, we connected the stepper motor to a stepper motor drive (Nudrive 4SX-211, National Instruments), which communicated with a computer (Power Macintosh G3, Macintosh, Mac OS 9.1) via a closed-loop motor-controller board (PCI-step-4CX, National Instruments). The computer ran a custom LabVIEW program (version 6.1, National Instruments) to conduct the experiment.

A B C

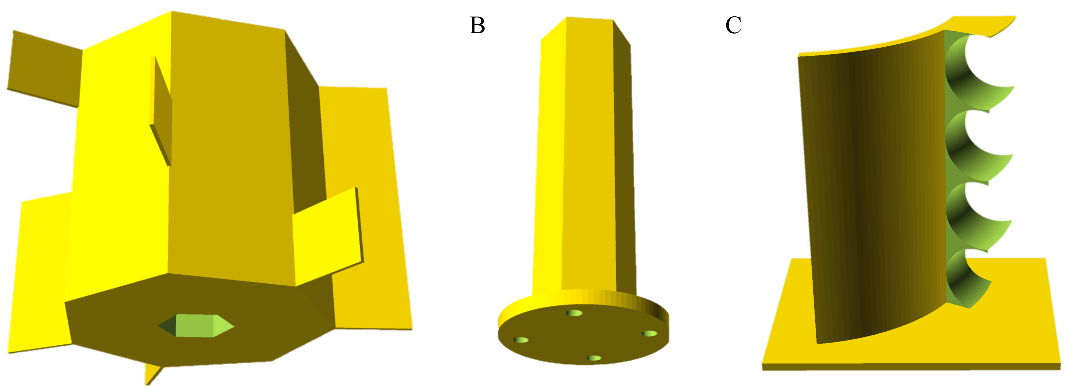

Figure 2. 3D Apparatus Design. ***A***, Discriminandum wheel containing the edges used in all 7 conditions. From left to right, D34, D2, D3, D4, and D234. ***B***, Post used to connect the discriminadum wheel to the stepper motor. ***C***, Support structure that held the participants' hand in place.

To fix the hand in a comfortable configuration to the discriminandum wheel and post, we 3D-printed a vertical support structure (Figure 2C). The hand-support structure was a 4-mm thick portion of a 126-mm diameter cylinder with a height of 110 mm and a 3-mm thick base 80 mm wide by 100 mm long. The structure had four protrusions of 28-mm diameter and 32.5-mm width, representing the segment of the finger overlying the proximal phalanx of each digit tested and an additional portion to support the pinky finger. To make the hand more comfortable, we molded Model Magic molding material (Crayola) onto the hand support structure. The hand-support structure was taped to the table approximately 10 mm from the edge of the discriminandum wheel and angled in such a way that the fingertips would align parallel to the edges of the discriminandum wheel. To ensure that the participants' digits were not touching the edges at any time that the stepper motor rotated the discriminadum wheel, we installed an infrared retro-reflective sensor (FE7B-RB6VG-M, Honeywell) parallel with the tested edge in order to monitor finger position. We used a magnetic plate with a magnetic arm to align the sensor when necessary.

2.4.3 Psychophysics. We devised a two-interval forced choice (2-IFC) task where participants discriminated the proximal/distal location of edges by slightly flexing the D2, D3, and D4 proximal interphalangeal joint (approximately 5 mm traveled by the finger) in order to contact the edge with the distal fingerpads. The digits that contacted the edge on a given experimental block were determined by the edge presented by the computer on that testing block. The two intervals of each trial presented edges at different proximal/distal locations and the participant was required to identify whether the edge in the second interval was more distal or proximal to the edge in the first interval. We defined the stimulus level as the proximal-distal distance (mm) between the two edges presented on a trial (Figure 3). We used a total of 24 stimulus levels. The minimal stimulus level was 0.1775 mm, which was determined by calculating the circumference ($2\pi r$) of the discriminandum wheel and dividing it by the number of steps the stepper motor took per revolution (2000 steps/revolution); the maximum stimulus level was 4.26 mm. The first trial tested started at 1.5 mm.

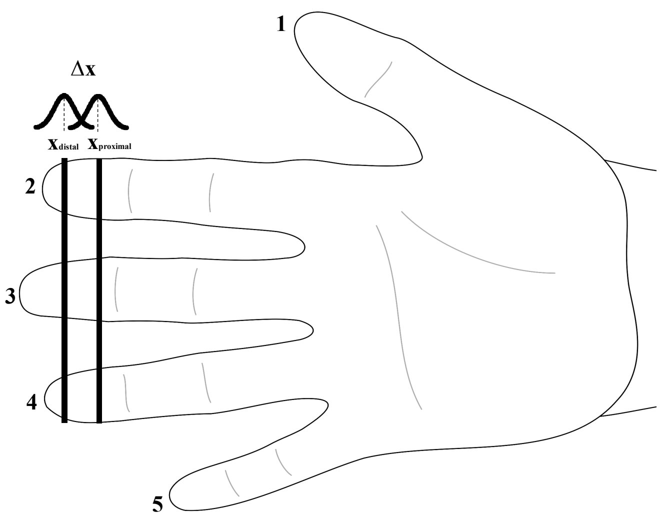

Figure 3. Measurement Distributions.
The computer program presented the participant with A distal and proximal stimulus. This figure uses D234 condition as an example.

We used a Bayesian adaptive method, known as the Ψ method (Kontsevich & Tyler, 1999), to efficiently estimate participants' psychometric functions. After each response, the algorithm updates the posterior probability of each function from a bank of many thousands of possible psychometric function shapes. The algorithm then uses an expected entropy procedure to choose the next stimulus, maximizing the information gain. We modeled each participant's psychometric function $P_c(\Delta x)$, the probability of a correct response as a function of stimulus, as a mixture of a cumulative normal (probit) function and lapse rate ε,

$$P_c(\Delta x) = 0.5\varepsilon + (1-\varepsilon)\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\Delta x/\sqrt{2}\sigma} exp\left(-\frac{y^2}{2}\right)dy, \tag{2.1}$$

where the σ-parameter is the threshold, corresponding to 76% correct response probability ($d' = 1$). The algorithm started with uniform prior probability distributions over a wide range of σ (0.1-5.0) and ε (0.01-0.08) values. The algorithm marginalizes the (σ, ε) joint posterior PDF in order to obtain the posterior PDF for σ. We read out the mode of the σ posterior PDF as the best estimate of the participant's σ value. We modified the Ψ algorithm by treating the lapse rate as a parameter of unknown value, as opposed to the value of 0.04 suggested by Kontsevich and Tyler, and by calculating a guessing Bayes factor after each trial:

$$BF = \frac{P(D|chance)}{P(D|(\sigma,\varepsilon)^*)} \tag{2.2}$$

This Bayes factor represents a likelihood ratio that compares the probability of the data under the hypothesis that the participant is guessing to the probability of the data under the hypothesis that the participant has a psychometric function. The numerator of the Bayes factor is the probability of the participant's data D, correct and incorrect responses at each of the stimulus levels, given that the participant is simply guessing (50% chance of guessing correctly) on all trials up to and including the current trial. The denominator is the probability of the data given the algorithm's best estimate of the participant's psychometric function. In most cases, the participant's Bayes factor rapidly approaches zero as the experiment progresses from trial-to-trial, indicating that the participant's performance conforms to a psychometric function. Note that all BFs were less than one.

2.4.4 Optimal cue-combination model. Adopting the Bayesian cue-combination paradigm (Trommershauser, Kording & Landy, 2011), we modelled the sensory measurement of the edge's location across D2, D3, and D4 as a sample drawn from a Gaussian centered about the actual location (Figure 3). According to the model, if the participant feels the edge x across all three digits, then the optimal combined percept (see Appendix B for derivation),

$$x' = \frac{\frac{x_2}{\sigma_2^2} + \frac{x_3}{\sigma_3^2} + \frac{x_4}{\sigma_4^2}}{\frac{1}{\sigma_2^2} + \frac{1}{\sigma_3^2} + \frac{1}{\sigma_4^2}}, \tag{2.3}$$

will on average simply equal x because the location of x on each digit is the same, $x_2 = x_3 = x_4$. We obtained the σ values, the stimulus levels that the participant could detect correctly 76% of the time, of each digit by testing each digit individually. We then calculated the variance of the optimal combined percept,

$$\sigma_{x'}^2 = \frac{1}{\frac{1}{\sigma_2^2} + \frac{1}{\sigma_3^2} + \frac{1}{\sigma_4^2}}, \tag{2.4}$$

for all multiple-digit combinations (D23, D24, D34, and D234). The square root of this variance is the model's prediction of the 76% threshold in the corresponding multiple-digit combination condition.

2.4.5 Procedure. The participants were seated sideways to a table with their right elbow supported by a foam pad and their hand gently gripping the hand-support structure. Each participant started the experiment by touching the D234 edge after having a dot drawn on D2 or D4, depending which one was the median length of the three, for the purpose of generating a 'home position', a highlighted line across all three digits, that was used to align the edges to the fingerpads between each block. The participants used a response device held in their left hand. We tested each participant individually.

On each trial, participants had to respond by pressing one of two buttons on the response device, indicating whether the second time they touched the edge was distal (towards the tips of their fingers) or proximal (towards their palm) relative to the first. On each touch, the participants received a signal (beep) to initiate movement of their fingers towards the edge and a signal to move their fingers back to the resting position. After each trial, participants received auditory feedback as to whether they were correct or incorrect.

The experiment consisted of seven conditions (D2, D3, D4, D23, D24, D34, and D234) (Figure 5) that were pseudorandomized, counterbalancing the single (D2, D3, and D4)- and triple (D234)-digit conditions. The participants repeated the seven conditions twice, for a total of 14 blocks. Each block consisted of 50 trials, taking an average of 6 minutes to complete, except for the first block, which had an additional 20 practice trials. Participants took a minimum 2-min break in-between each block and a 5-min break halfway through the experiment.

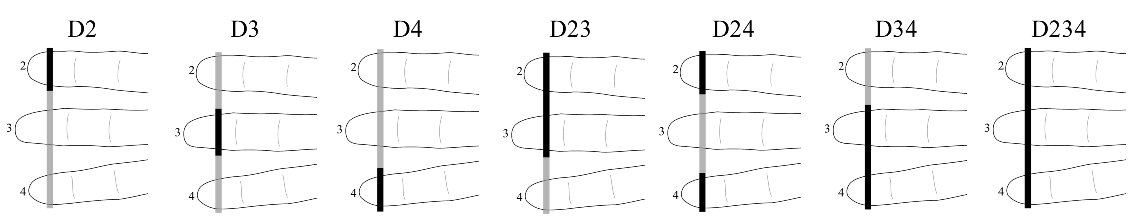

Figure 4. Stimuli. The experimental conditions consisted of all three individual fingers and their four combinations. The line across the three digits represents a stimulus, where the black region is the digit(s) tested and the gray region is the digit(s) not tested. From the testing, we obtained sigma σ values for individual digits, as well as multiple digits.

2.4.6 Statistical Analysis. We performed repeated measures analysis of variance (RM-ANOVA) using SPSS Statistics v19 (IBM) for Macintosh, with an ∝-level of 0.05. The RM-ANOVA models were type III sum-of-squares, testing for main effects of all factors.

If the analysis failed Mauchly's test of sphericity, we used the Greenhouse-Geisser correction. We used Bonferroni adjustment for all main effect pairwise comparisons. The mode of the σ parameter posterior PDF (participants' 76%-correct threshold) was the dependent measure used in all statistical analyses. We analyzed both the average of the two replicates and the best (i.e., lowest) σ-mode of the two replicates.

To determine if there was a significant difference between the measured and optimal combined σ values, we computed 2 (σ variant: Measured vs. Optimal) x 4 (Multiple-Digit Combinations: D23, D24, D34, vs. D234) RM-ANOVAs. To determine whether the multiple-digit σ values were less than the respective single-digit σ values, an assumption of the optimal cue-combination formula, we computed three one-way RM-ANOVA.

2.4 Results

Figure 6 shows the experimental data for the average and best of the two replicates, where D2, D3, and D4 represent the single-digit σ that were used to calculate the respective multiple-digit optimal cue combination. The figure indicates that the measured multiple-digit means were higher than the predicted optimal multiple-digit means. Additionally, the multiple-digit measured means were not generally lower than the single-digit means. These observations held for both the average of the replicates (Fig. 6A) and the best replicate (Fig. 6B). Figure 7 show the within-participant difference scores (measured minus optimal predicted thresholds) for the multiple-digit combinations, once again indicating that the measured multiple-digit thresholds were higher than the predicted optimal multiple-digit thresholds. Again, this trend occurred in both the average and the best replicates. Collectively, these graphs strongly suggest that the predictions of the optimal cue combination formulae were not met. The statistical analyses (below) confirmed this impression.

2.5.1 Replicate Average σ. To determine whether participants combines cues from multiple digits optimally, we conducted a 2 (σ variant: Measured vs. Optimal) x 4 (Multiple-Digit Combinations: D23, D24, D34, vs. D234) RM-ANOVA. This analysis showed that the measured σ values were significantly greater than the optimal. There was a significant main effect of σ variant ($F_{(1,23)} = 14.063, p < .001$) (Figure 6A) and a significant main effect of Digit Combination (F…) with no significant interaction.

A predication of the optimal cue combination model is that the combined sigma values will be less than any of the single-digit sigma values. However, the one-way RM-ANOVA on the σ values of D2, D23, D24, and D234 showed a significant difference ($F_{(3,69)} = 4.650, p = .005$), where D2 was significantly less than D23 and D234 ($p = .06$ and $p = .049$, respectively). The one-way RM-ANOVA for D3, D23, D34, and D234 was not significant. The one-way RM-ANOVA for D4, D24, D34, and D234 measured σ values

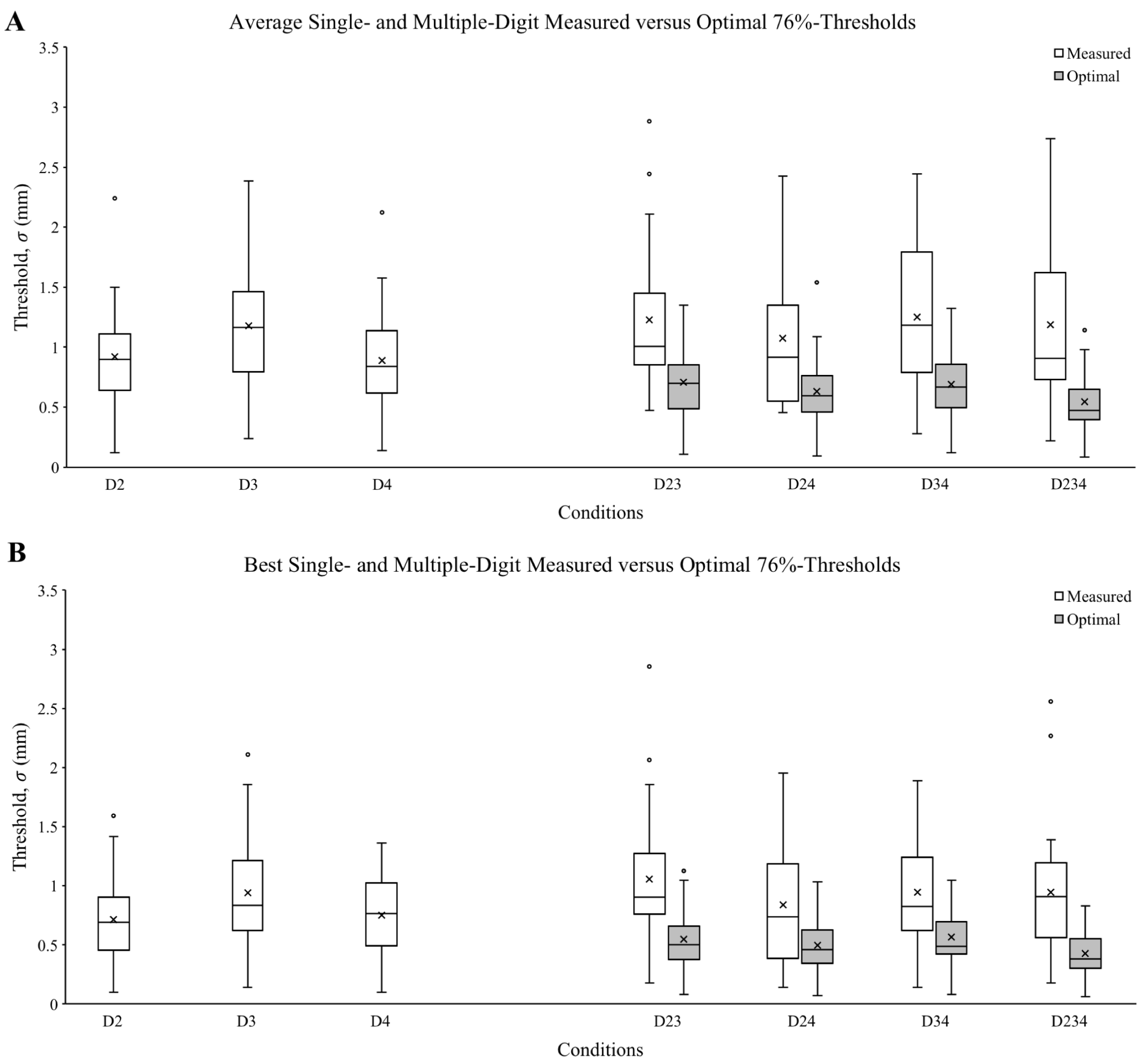

Figure 5. 76%-Thresholds. ***A***, 76%-thresholds for the average of the two replicates. ***B***, 76%-thresholds for the best of the two replicates. The x-axes represent the 76%-threshold in mm (i.e. the σ value) and the y-axes represent the experimental conditions. The error bars represent the outer quadrants (quadrants one and four) and the boxes represent quadrants three and four. The x's within the boxes are the mean and the line represents the median. Measured σ values are in white and the calculated optimal σ values are in grey.

was significant ($F_{(3,69)} = 6.355, p = .001$), where D4 was significantly less than D34 and D234 ($p = .002$ and $p = .015$, respectively). These analyses indicate, contrary to the prediction, that the index and ring digits performed better when tested alone than in combination with other digits.

2.5.2 Replicate Best σ. Next, we repeated the preceding analyses using the best (i.e. the lowest) sigma value of the two replicates. The 2X4 RM-ANOVA showed that the measured σ values were significantly greater than the optimal σ values ($F_{(1,23)} = 56.275, p < .001$) (Figure 6B). The one-way RM-ANOVA for D2, D23, D24, and D234 measured σ values was significant ($F_{(3,69)} = 4.458, p = .006$), where D2 was significantly less than D23 ($p = .01$). The one-way RM-ANOVA for D3, D23, D34, and D234 was not significant. The one-way RM-ANOVA for D4, D24, D34, and D234 was not significant. As in the previous section, these analyses do not provide evidence to support optimal cue combination.

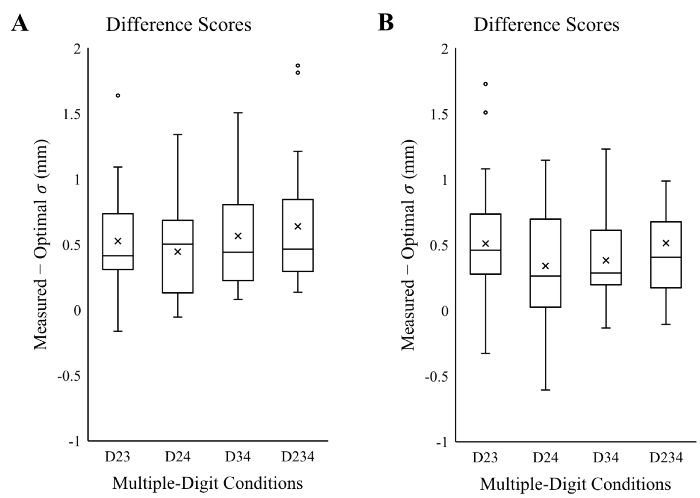

Figure 6. Difference Scores. Average and best of replicates 76%-thresholds for multiple-digit combinations.

2.5 Discussion

Here, we tested our hypothesis about whether cues from multiple digits are combined optimally using a 2IFC discrimination task with a 1-mm thick edge stimulus. Overall, the empirical data obtained in this study do not provide evidence for optimal cue combination between the tested digits. The statistical analysis showed a significant difference between the empirically measured combined σ values and the ones calculated using the cue combination formula, suggesting that we failed to measure cue combination that was equivalent to optimal. In contrast to the prediction of optimal cue combination that the σ values of multiple-digits should be less than those of the individual digits, the results of this study showed the opposite trend.

The results of this experiment contradict previous research where optimal cue combination was observed in the majority of the studies outlined in Chapter 1. However, the results of this study are consistent with the work done by Rosas and colleagues (2005). That study looked at slant and texture cues between the tactile and visual modalities, respectively, to determine whether these cues were integrated according to the optimal cue combination model. The experimental design was similar to the one use in this experiment in that the participants had to respond by pressing one of two buttons that corresponded to the more slanted plane, where in our experiment the participants indicated which of the two stimuli was more distal. The study conducted by Rosas and colleagues (2005) was almost a replicate of a study conducted by Ernst and Banks (2002), but there was one key difference: Ernst and Banks (2002) used cue conflicts. Neither the current study nor that of Rosas and colleagues (2005) introduced cue conflicts to the participants. It is possible that, with the use of a cue conflict paradigm, our study would have revealed evidence of optimal cue combination. However, further research is needed to confirm this hypothesis.

With the current results, we can use the theoretical framework described in the previous chapters to speculate what experimental and perceptual properties could have given rise to the data observed in this experiment. Let us consider the hypothesis that discrimination of spatial cues across multiple digits is not optimal. If that is the case, then that implies that multiple-digit cue combination violates one of the four assumptions of the ideal model: 1. The cues arise from a single source; 2. The cues from different digits provide independent information; 3. The measurement distributions are Gaussian; and 4. The observer knows the measurement σ.

It is also possible that cue combination between digits is optimal, but we just failed to observe it under the current experimental paradigm. There are a number of ways we can modify the current experimental design to reduce the cue uncertainty. Making the edges thinner would result in a sharper measurement distribution and in turn, the observer would produce a sharper likelihood function. It could also be the case that in order for optimal cue combination to occur between the digits, the stimulus has to contact each digit at the same relative area. Since we observed that D3 was significantly worse than the other two digits, it could have been due to where we tested on the digit (more proximally on the fingerpad than the other two digits).

Taken together, the results of this experiment warrant further investigation into determining whether cues from multiple digits can be combined in an optimal manner. In the next chapter, we will take a detailed look at the assumptions of the ideal observer model and outline potential experimental design flaws that could have contributed to the results. We will also outline future directions with respect to modifications that could be made to the current experimental paradigm.

2.6 References

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

Alais D, Burr D (2004) No direction-specific bimodal facilitation for audiovisual motion detection. Brain Res Cogn Brain Res 19:185-194.

Battaglia PW, Jacobs RA, Aslin RN (2003) Bayesian integration of visual and auditory signals for spatial localization. JOSA A 20:1391-1397.

Bergmann Tiest WM, Kappers A (2009) Cues for Haptic Perception of Compliance. IEEE Trans Haptics 2:189-199.

Burr D, Banks MS, Morrone MC (2009) Auditory dominance over vision in the perception of interval duration. Exp Brain Res 198:49-57.

C. Craig JMK, James (1998) Factors affecting tactile spatial acuity. Somatosens Mot Res 15:29-45.
Craig JC (1999) Grating orientation as a measure of tactile spatial acuity. Somatosens Mot Res 16:197-206.
Craig JC (2003) The effect of hand position and pattern motion on temporal order judgments. Percept Psychophys65:779-788.
Craig JC, Qian X (1997) Tactile pattern perception by two fingers: temporal interference and response competition. Perception & psychophysics 59:252-265.
Debats NB, van de Langenberg RW, Kingma I, Smeets JB, Beek PJ (2010) Exploratory movements determine cue weighting in haptic length perception of handheld rods. J Neurophysiol 104:2821-2830.
Drewing K, Wiecki TV, Ernst MO (2008) Material properties determine how force and position signals combine in haptic shape perception. Acta Psychol (Amst) 128:264-273.
Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. Nature 415:429.
Gepshtein S, Banks MS (2003) Viewing Geometry Determines How Vision and Haptics Combine in Size Perception. Current Biology 13:483-488.
Hartcher-O'Brien J, Di Luca M, Ernst MO (2014) The duration of uncertain times: audiovisual information about intervals is integrated in a statistically optimal fashion. PLoS One 9:e89339.
Hillis JM, Watt SJ, Landy MS, Banks MS (2004) Slant from texture and disparity cues: optimal cue combination. J Vis 4:967-992.
Jacobs RA (1999) Optimal integration of texture and motion cues to depth. Vision research 39:3621-3629.
Knill DC, Saunders JA (2003) Do humans optimally integrate stereo and texture information for judgments of surface slant? Vision Research 43:2539-2558.
Kuschel M, Di Luca M, Buss M, Klatzky RL (2010) Combination and Integration in the Perception of Visual-Haptic Compliance Information. IEEE Trans Haptics 3:234-244.
Landy MS, Kojima H (2001) Ideal cue combination for localizing texture-defined edges. JOSA A 18:2307-2320.
Lederman SJ, Klatzky RL (2009) Haptic perception: a tutorial. Atten Percept Psychophys 71:1439-1459.
Peters RM, Hackeman E, Goldreich D (2009) Diminutive digits discern delicate details: fingertip size and the sex difference in tactile spatial acuity. J Neurosci 29:15756-15761.
Read JC (2012) Visual perception: understanding visual cues to depth. Curr Biol 22:R163-165.
Rosas P, Wagemans J, Ernst MO, Wichmann FA (2005) Texture and haptic cues in slant discrimination: reliability-based cue weighting without statistically optimal cue combination. JOSA A 22:801-809.
Takahashi C, Diedrichsen J, Watt SJ (2009) Integration of vision and haptics during tool use. J Vis 9:3 1-13.

Trommershauser J, Kording K, Landy MS (2011) Sensory cue integration: Oxford University Press.
Wismeijer DA, Erkelens CJ, van Ee R, Wexler M (2010) Depth cue combination in spontaneous eye movements. J Vis 10:25.
Wong M, Peters RM, Goldreich D (2013) A physical constraint on perceptual learning: tactile spatial acuity improves with training to a limit set by finger size. J Neurosci 33:9345-9352.

CHAPTER 3

GENERAL DISCUSSION

The Bayesian perceptual framework describes mathematically how an optimal perceptual inference can be obtained on the basis of multiple sources of uncertain sensory information, a process known as cue combination. An important feature of Bayesian cue combination is that every sources of stimulus-evoked input, no matter how noisy, provides relevant information; therefore, an inference based on more than one sensory cue will be more accurate than an inference based on any of the individual cues alone. The Bayesian cue combination framework has been applied extensively in multisensory studies and in visual studies, but rarely in haptic studies. The goal of the research reported in this thesis was to determine whether or to what extent human participants could achieve optimal sensory integration of inputs from the fingers. We tested 24 participants with a 2-IFC edge-location discrimination experiment involving the index, middle, and ring fingers individually and in each of four possible combinations. The Bayesian cue combination formulas predicted that our participants would perform more accurately with fingers in combination than with any one finger alone. Interesting, the experimental results did not confirm this prediction. Here we discuss some possible reasons for this surprising negative finding.

In order to determine whether cues from individual digits are estimated by participants according to the optimal observer model, we presented 1-mm thick edges across the width of each digit tested. A Bayesian adaptive algorithm adjusted the stimulus level Δx, the difference in mm between the distal $x_{\uparrow}$ and proximal $x_{\downarrow}$ edge locations. The participant responded by indicating which interval represented the more distal location. For each of the seven experimental conditions (D2, D3, D4, D23, D24, D34, and D234), the algorithm began with $\Delta x = 1.5\ mm$, choosing the next Δx from a range of $0.1735\ mm$ to $4.26\ mm$ based on expected entropy minimization, until the last trial where the algorithm estimated the observer's psychometric function CDF and the 76%-threshold value. The 76%-threshold of a CDF that results from the difference between measurement distribution of random variables (RVs) $X_{\uparrow}$ and $X_{\downarrow}$ is equal to the standard deviation of the measurement distribution when the sensitivity index $d' = 1$. This Δx value on the psychometric function at 76% correct is the σ value of the participant's spatial measurement distributions. Another analytically relevant %-threshold value is the Δx at 84%-correct, which corresponds to the standard deviation of the psychometric function CDF (i.e., $\sqrt{2}\sigma$). In fact, it does not matter which %-correct is chosen to represent the threshold. The cue-combination formulas hold for any sigma value that is a constant multiple of the measurement distribution sigma, because the constant factor cancels out in the formulas.

An experiment is constructed under certain conditions that the researcher tries to either control or measure in order to have a more complete picture of the phenomena under investigation. One such phenomenon of interest to computational neuroscientists is an analytical model with the capacity to predict/accurately estimate behaviour, in general terms, and more specifically, to predict behaviour in response to specific sensory cues where a sensory cue is any environmental/external perturbation of the receptors that elicits a change in perception in the observer via transduction of the information carried by the cue at the time of 'contact' with a sensory receptor. I use contact to represent anytime an observer's receptors are stimulated by any externally arising change in neural signaling that elicits a response in the observer. One such method with the capacity to predict human behaviour is the ideal observer model used as the theoretical framework in our psychophysics experiment. The ideal observer model makes four main assumptions that deserve a closer look, as the violation of any of these assumptions might have led to the negative findings of the current experiment.

1. The cues come from a single source: $\boldsymbol{p(x_2, x_3, x_4 | s)}$

The model assumes that each measurement comes from the same source – i.e., that the position of the edge in each interval of the 2IFC trial was the same for all digits stimulated. In other words, the model assumes that the mean of the measurement distribution from which the observed data were sampled is the same on all three digits. If this assumption is violated, either by the experimental conditions or by the observer's familiarity with the stimulus in the experimental context, then the predictions of the cue combination formula will not hold.

One of our experimental conditions, the D24 condition, indeed violated the single-source assumption. The D24 condition is actually two separate edges with a gap between them; therefore, when the participants made contact with the D24 stimulus, in the absence of a stimulation to D3, they might have concluded that in fact they should not combine that information according to the cue combination formula. It is possible, therefore, that the apparently sub-optimal performance of our participants in the D24 condition owes to this circumstance.

2. The cues are conditionally independent: $\boldsymbol{p(x_2, x_3, x_4 | s) = p(x_2 | s) p(x_3 | s) p(x_4 | s)}$.

A second assumption made by the ideal observer model is that the cues are conditionally independent, given the source. In other words, the ideal observer model assumes that each measurement is an independent sample drawn from the corresponding finger's Gaussian measurements distribution. If the samples are not conditionally independent, but rather correlated (as could happen for example if sensory neurons from the three digits are synaptically connected to one another), then the assumption of conditional independence is violated and the predictions of the cue combination formula will not hold up empirically.

3. $\boldsymbol{X_\uparrow}$ and $\boldsymbol{X_\downarrow}$ RVs are $\boldsymbol{\approx \mathcal{N}(x, \sigma)}$

The ideal observer model assumes that the cues are Normally distributed. A Gaussian measurement distribution may be a good approximation, but it is an idealization of a much more complex underlying neurophysiology, in which the true measurements are not single Gaussian samples but rather action potential counts from numerous responding neurons that are subject to cortical neural variability (i.e., Poisson noise). For this reason alone, some mismatch is to be expected between the exact predictions of the ideal observer model and the empirical performance of human participants.

4. The observer knows its measurement σ.

Implicit to the use of the cue combination formula for optimal inference is the assumption that the observer has access to the correct sigma values. If the sigma values entered into the cue combination formula do not equal the actual sigma values of the measurement distributions for the stimuli and fingers involved in the task, then the inference will be sub-optimal. An intriguing possibility is that the suboptimal performance of our participants may be due to their lack of familiarity with the particular tactile-edge stimulus and task that we constructed for them. In this case, it could be helpful to provide the participant more experience with the task. Training participants extensively on the task might improve their performance by allowing them to learn the correct sigma values.

3.1 Future Direction

While the results of this thesis do not provide conclusive evidence for or against optimal cue combination between the digits tested, the results, along with the discussion of potential theoretical and empirical sources of error, do provide a concrete basis for the direction of future research. Here, we focus on experimental modification that might improve the odds of observing optimal cue combination between digits as well as explore the idea that cues from multiple digits might not be optimal, providing alternative models that might better explain how cues from multiple digits are integrated by the nervous system.

As mentioned above, one obvious extension of our current experiment would be to repeat the current protocol over multiple days of training with feedback. Such an experiment would allow us to determine whether participants can improve their ability to combine cues with practice.

An intriguing modification to the design used in this study would be a cue-conflict task, in which we would surreptitiously offset the location of either $x_{\uparrow}$ or $x_{\downarrow}$ between digits in one of the two intervals. Using the D23 condition and stimulus level $\Delta x = 1.5\ mm$ as an example, if we present D2 with $x_{\uparrow} = 0\ mm$, then we would present D3 with a small offset from 0 mm, such as $x_{\uparrow} = 0.05\ mm$ in one interval (the conflict interval), and then present $x_{\downarrow} = 1.5\ mm$ to both D2 and D3 in the second interval (the non-conflict interval). We could then use the Bayesian adaptive algorithm to adaptively adjust Δx in order to estimate the participant's point of subjective equality (PSE), the Δx at which the participant perceives the nonconflict and conflict locations to be identical. The optimal cue combination formula for x' (see Chapter 2) predicts the influence of the conflict offset (0.05 mm in this example) on the PSE, and the participant's data would be compared to that prediction.

Another modification to our design would be to use curved edges rather than straight ones. The potential advantage of a curved edge is that it could be made to contact all three digits on the same relative location, as referenced to the center of the fingerpad, for example. This might eliminate the potential problem that we encountered of higher-than-expected thresholds on D3.

A final modification to our design would be to allow multiple or continuous rather than binary response choices. For instance, we could introduce a response device that would allow the participant to point to the perceived location of a single edge. In theory, this method allows a straightforward and fine-resolution way to gauge the participant's perception of edge location, without the use of a 2-IFC discrimination experiment. In practice, however, such a method might be subject to motor variability that could make the data more difficult to interpret.

It's possible that we won't observe optimal cue combination between digits even with all the modification. In that case, we ought to consider using a different model to predict participant's behaviour on this task. If it is the case that the observer cannot perceive the cues from each digit as a single source, due to the nervous system or the experimental design, we could use a model that predicts performance when the cues indeed arise from different sources. On the other hand, the observer could be interpreting the source differently, depending on the trial. In this case, we could use a mixture model, such as the causal inference model derived by Kording et al. (2007).

3.2 Conclusion

Overall, the theoretical framework outlined and empirical research conducted in this thesis provide some insight into cue combination between multiple digits. Because we did not observe optimal cue combination, we discuss several theoretical assumptions that could have been violated and propose modification to the current experimental paradigm that might sharpen our understanding of whether cue combination between digits is truly sub-optimal, in which case we suggest an alternative model that might better explain how cues from multiple digits are combined by the nervous system.

3.3 References

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

APPENDIX A

BAYES' THEOREM DERIVATION

Let $P(A|B)$ denote the conditional probability of A and B. Then for ant two events A and B with $P(A) > 0$ and $P(B) > 0$, the conditional probability of A given that B has occurred is defined by:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{A.1}$$

And the conditional probability of B given that A has occurred is define by:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \tag{A.2}$$

By the product rule for probabilities, it follows that:

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A) \tag{A.3}$$

Dividing both sides by $P(A)$ results in the following:

$$\frac{P(A \cap B)}{P(A)} = \frac{P(A|B)P(B)}{P(A)} = P(B|A) \tag{A.4}$$

Therefore:

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} \tag{A.5}$$

This equation is known as Bayes' Theorem. We can now partition event B and derive the extended form of Bayes' Theorem.

Let $B_1, \ldots, B_N$ be mutually exclusive and exhaustive events. Then for any other event A:

$$P(A) = P(A|B_1)P(B_1) + \cdots + P(A|B_N)P(B_N) = \sum_{k=1}^{N} P(A|B_k)P(B_k) \tag{A.6}$$

It follows that for any B_i:

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{k=1}^{N} P(A|B_k)P(B_k)} \tag{A.7}$$

Now let event A represent some data D and event B represent some hypothesis H, then it follows that:

$$P(H_i|D) = \frac{P(D|H_i)P(H_i)}{\sum_{k=1}^{N} P(D|H_k)P(H_k)} \tag{A.8}$$

This form of Bayes' Theorem allows us to compute the posterior probability of each hypothesis given some data.

APPENDIX B

CUE COMBINATION FORMULA DERIVATION USING CALCULUS

The source of a stimulus at position s produces a measurement drawn from a Gaussian distribution centered on s. Now consider three Likelihood Functions (LFs), resulting from index (D2), middle (D3), and ring (D4) finger measurements x_2, x_3, and x_4, respectively, where each LF is also a Gaussian with its own mean μ and variance σ^2. Then the combined LF is

$$p(x_2, x_3, x_4|s) = p(x_2|x_3, x_4, s)p(x_3|x_4, s)p(x_4|s). \tag{B.1}$$

Under the assumption that the measurements x_2, x_3, and x_4 are conditionally independent, given the stimulus, then,

$$p(x_2, x_3, x_4|s) = p(x_2|s)p(x_3|s)p(x_4|s). \tag{B.2}$$

Given that the LFs are Normally distributed $\mathcal{N}(\mu, \sigma^2)$, then,

$$p(x_2, x_3, x_4|s) = \frac{1}{\sigma_2\sqrt{2\pi}} e^{-\frac{(s-x_2)^2}{2\sigma_2^2}} \frac{1}{\sigma_3\sqrt{2\pi}} e^{-\frac{(s-x_3)^2}{2\sigma_3^2}} \frac{1}{\sigma_4\sqrt{2\pi}} e^{-\frac{(s-x_4)^2}{2\sigma_4^2}}. \tag{B.3}$$

Taking the natural log of the likelihood, we find that:

$$\ln p(x_2, x_3, x_4|s) = C - \ln\sigma_2^2 - \frac{(s-x_2)^2}{2\sigma_2^2} - \ln\sigma_3^2 - \frac{(s-x_3)^2}{2\sigma_3^2} - \ln\sigma_4^2 - \frac{(s-x_4)^2}{2\sigma_4^2} \tag{B.4}$$

We then differentiate with respect to s, finding that:

$$\frac{d}{ds}\ln p(x_2, x_3, x_4|s) = -\frac{s-x_2}{\sigma_2^2} - \frac{s-x_3}{\sigma_3^2} - \frac{s-x_3}{\sigma_3^2} \tag{B.5}$$

To determine the maximum likelihood estimate (MLE), $\hat{s}$, which is the mode at which $\frac{d}{ds}\ln p(x_2, x_3, x_4|s) = 0$, we get:

$$0 = \frac{\hat{s}-x_2}{\sigma_2^2} + \frac{\hat{s}-x_3}{\sigma_3^2} + \frac{\hat{s}-x_4}{\sigma_4^2} \tag{B.6}$$

It follows that:

$$\hat{s}\left(\frac{1}{\sigma_2^2}+\frac{1}{\sigma_3^2}+\frac{1}{\sigma_4^2}\right)=\frac{x_2}{\sigma_2^2}+\frac{x_3}{\sigma_3^2}+\frac{x_4}{\sigma_4^2} \tag{B.7}$$

Isolating $\hat{s}$, the perceived position of the stimulus across all three digits, we get:

$$\hat{s}=\frac{\frac{x_2}{\sigma_2^2}+\frac{x_3}{\sigma_3^2}+\frac{x_4}{\sigma_4^2}}{\frac{1}{\sigma_2^2}+\frac{1}{\sigma_3^2}+\frac{1}{\sigma_4^2}} \tag{B.8}$$

Note that if the variances are equal, $\hat{s}$ will be the simple arithmetic average of the measurements. If one of the variances is really large, $\hat{s}$ will tend towards the weighted average of the other two measurements.

Continuing to differentiate from $\ln p(x_2, x_3, x_4|s)$, we take the second derivative to determine the combined variance:

$$\frac{d^2}{ds^2}\ln p(x_2, x_3, x_4|s) = -\frac{1}{\sigma_2^2}-\frac{1}{\sigma_3^2}-\frac{1}{\sigma_4^2} \tag{B.9}$$

Assuming that the product of the individual LFs is a Gaussian, it will have variance:

$$\sigma^2=\frac{1}{-\frac{d^2}{ds^2}\ln p(x|s)} \tag{B.10}$$

$$\sigma_{2,3,4}^2=\frac{1}{\frac{1}{\sigma_2^2}+\frac{1}{\sigma_3^2}+\frac{1}{\sigma_4^2}} \tag{B.11}$$