

---

Electronic Theses and Dissertations, 2004-2019

---

2011

## Perceptual Judgment The Impact Of Image Complexity And Training Method On Category Learning

Michael T. Curtis  
*University of Central Florida*

 Part of the [Psychology Commons](#)

Find similar works at: <https://stars.library.ucf.edu/etd>

University of Central Florida Libraries <http://library.ucf.edu>

This Doctoral Dissertation (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations, 2004-2019 by an authorized administrator of STARS. For more information, please contact [STARS@ucf.edu](mailto:STARS@ucf.edu).

---

### STARS Citation

Curtis, Michael T., "Perceptual Judgment The Impact Of Image Complexity And Training Method On Category Learning" (2011). *Electronic Theses and Dissertations, 2004-2019*. 1835.

<https://stars.library.ucf.edu/etd/1835>

PERCEPTUAL JUDGMENT: THE IMPACT OF IMAGE COMPLEXITY AND  
TRAINING METHOD ON CATEGORY LEARNING

by

MICHAEL T. CURTIS  
B.A. University of Kentucky, 2001  
M.A. University of Central Florida, 2008

A dissertation submitted in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy  
in the Department of Psychology, Applied Experimental Human Factors,  
in the College of Sciences  
at the University of Central Florida  
Orlando, Florida

Fall Term  
2011

Major Professor: Florian G. Jentsch

© 2011 Michael T. Curtis

## **ABSTRACT**

The goal of this dissertation was to bridge the gap between perceptual learning theory and training application. Visual perceptual skill has been a vexing topic in training science for decades. In complex task domains, from aviation to medicine, visual perception is critical to task success. Despite this, little, if any, emphasis is dedicated to developing perceptual skills through training. Much of this may be attributed to the perceived inefficiency of perceptual training. Recent applied research in perceptual training with discrimination training, however, holds promise for improved perceptual training efficiency. As with all applied research, it is important to root application in solid theoretical bases. In perceptual learning, the challenge is connecting the basic science to more complex task environments. Using a common aviation task as an applied context, participants were assigned to a perceptual training condition based on variation of image complexity and training type. Following the training, participants were tested for transfer of skill. This was intended to help to ground a potentially useful method of perceptual training in a model category learning, while offering qualitative testing of model fit in increasingly complex visual environments. Two hundred and thirty-one participants completed the computer based training module. Results indicate that predictions of a model of category learning largely extend into more complex training stimuli, suggesting utility of the model in more applied contexts. Although both training method conditions showed improvement across training blocks, the discrimination training condition did not transfer to the post training transfer tasks. Lack of adequate contextual information related to the transfer task in training was attributed to this outcome. Further analysis of the exposure training condition showed that

individuals training with simple stimuli performed as well as individuals training on more complex stimuli in a complex transfer task. On the other hand, individuals in the more complex training conditions were less accurate when presented with a simpler representation of the task in transfer. This suggests training benefit to isolating essential task cues from irrelevant information in perceptual judgment tasks. In all, the study provided an informative look at both the theory and application associated with perceptual category learning. Ultimately, this research can help inform future research and training development in domains where perceptual judgment is critical for success.

This dissertation is dedicated to the memory of my father, Stephen Curtis.

## ACKNOWLEDGMENTS

I share the accomplishment of completing my graduate education with a varied cast to whom I hold great veneration for the wisdom, patience, and support they imparted upon me.

First and foremost, I would like to express my gratitude to Dr. Florian Jentsch, my committee chair, advisor, mentor, and comrade. Thank you for providing me room to grow, opportunities to achieve, and the trust to lead. Your generosity, intelligence, and support have been instrumental to my development as a scientist. Above all else, you have shown me the strength of resolve in the face of adversity. I would also like to extend words of appreciation to my committee, Dr. Ladislau Boloni, Dr. Eduardo Salas, and Dr. James Szalma, for offering a spectrum of viewpoints, and valuable feedback which helped to strengthen this document and me as a researcher.

Second, I would like to acknowledge my colleagues at the Team Performance Laboratory whom I have worked with, learned from, and respect tremendously. Though too numerous to name; each were integral to my development. I will forever relish our spirited collaborations, lab unity, and the lasting friendships which I will take with me. I especially owe a debt of gratitude to Ron Swanson, Crystal Maraj, Monica Ritman, Jose Quevedo, and Javier Rivera who dedicated innumerable hours of their time over the life of the project which spawned this dissertation. I think it is safe to put away the Mike-face dartboards now.

Third, I would like to acknowledge the contribution my wife, Debbie, had on this journey. We rode the roller coaster of graduate school together, and it gives me great pride to

complete the journey together. Her undeterred patience, encouragement, and advice kept me from derailing numerous times. She reminded me that a smile and a laugh are more productive than yelling at inanimate objects any day of the week. For your continued confidence in my ability, I share this accomplishment with you, Pistolorada.

Special recognition goes to my mother Angela and father Stephen who provided unquestioned support for my crazy idea to go back to school. Although the cosmos had other plans for you, Dad, I will never forget the life affirming discussion we shared on the way to my campus visit in Orlando. Once in school, credit is also due to my brother Ryan who provided timely diversions to help me escape the relentless stress of graduate school.

Every journey begins with a single step, and for that first step I am indebted to my brother Phil. His idea for me to visit him on the west coast in the Summer of 2003 was the unforeseen catalyst for this journey. It is hard to believe that a ticket to a music festival and lazy afternoon on the Russian River would change my life. Special thanks goes to Brad, Christa, Chuck, Jane, Will, Tyler, and the rest of the San Francisco crew who have no idea what impact their words had on my future. They planted a seed in my mind that the world is bigger than I know. For that I am forever grateful.



## TABLE OF CONTENTS

LIST OF FIGURES .....	xii
LIST OF TABLES .....	xiii
CHAPTER 1: INTRODUCTION .....	1
Statement of Problem.....	1
Purpose of Study .....	3
CHAPTER 2: LITERATURE REVIEW .....	5
Perceptual Learning .....	7
Mechanisms of Perceptual Learning.....	8
Attention Weighting.....	8
Imprinting .....	9
Differentiation.....	10
Unitization.....	11
Summary .....	12
Perceptual Judgment .....	13
Category Learning .....	14
Models of Category Learning .....	16
Exemplar Models of Category Learning.....	17
Rule-Based Models of Category Learning .....	18
The ATRIUM Model .....	20
Effect of Stimulus Complexity on Model Fit .....	25
Perceptual Training.....	29
Perceptual Skill .....	30
Perceptual Training Method .....	32

Effect of Training Method on Category Learning .....	36
Moderating Effect of Training Method on Training Stimulus Complexity.....	39
Chapter Summary .....	40
CHAPTER 3: MATERIALS AND METHODS .....	43
Participants.....	43
Design .....	43
Experimental Task .....	44
Category Labels .....	47
Apparatus .....	48
Administration .....	48
Stimuli.....	49
Training.....	50
Training Method .....	50
Image Complexity.....	51
Control Group .....	56
Summary of Training.....	56
Measures .....	57
Demographics .....	57
Spatial Orientation .....	57
Skill Transfer .....	57
Procedure .....	58
CHAPTER 4: RESULTS.....	59
Descriptive Data.....	59
Hypothesis 1a-d: Effects of Stimulus Complexity on Model Fit.....	62

Hypothesis 1a: As training stimulus complexity increases, participant reliance on rule-based response strategies will increase such that overgeneralization will be evident in training for more trials than in simple stimulus conditions.....	62
Hypothesis 1b: As training stimulus complexity increases, participant reliance on rule-based response strategies will increase such that a lower proportion of exception training stimuli will be labeled correctly.....	65
Hypothesis 1c: As training stimulus complexity increases, participant reliance on rule-based response strategies will increase such that less overall exception responses will be given in the transfer task. ....	68
Hypothesis 1d: As complexity increases greater reliance on rule-based processes will impact dimensional attention in the transfer task such that a higher proportion of exception responses will occur on the primary (rule) dimension of exceptions than on the secondary (exception) dimension. ....	69
Hypothesis 2a-b: Effect of Training Method on Category Learning.....	72
Hypothesis 2a: By increasing the number of images per exposure, the discrimination training will result in reduced time to reach training proficiency. ....	72
Hypothesis 2b: Two image discrimination training elicits explicit practice with both exemplar and rule-based categorization strategies such that overall accuracy will be better on the transfer task than in the single exposure training condition. ....	74
Hypothesis 3a-b: Moderating Effect of Training Method on Training Stimulus Complexity.....	75
CHAPTER 5: DISCUSSION.....	78
Hypothesis Discussion.....	79
Hypothesis 1.....	79
Hypothesis 2.....	83
Hypothesis 3.....	84
Theoretical Implications .....	84
Model Fit.....	85
Training Type.....	86
Transfer Performance Differences in Exposure Training.....	88

Practical Implications.....	89
Limitations and Future Research .....	91
Conclusion .....	95
APPENDIX A: IRB APPROVAL LETTER .....	97
APPENDIX B: BIOGRAPHICAL DATA QUESTIONNAIRE .....	99
APPENDIX C: SPATIAL ORIENTATION MEASURE .....	102
APPENDIX D: TRAINING PERFORMANCE PLOTS.....	108
LIST OF REFERENCES .....	110

## LIST OF FIGURES

Figure 1. Graphical representation of experimental comparisons. ....	42
Figure 2. Sample static visual approach image. Generated using Microsoft Flight Simulator: X (Microsoft 2006). ....	45
Figure 3. Geometric properties of glideslope. ....	46
Figure 4. Precision Approach Path Indicator lighting configuration. ....	47
Figure 5. Percentage of overgeneralized responses for training Block 1, 14,15,and 16 by training image complexity condition for the exposure training condition. ....	64
Figure 6. Percentage of correct exception responses across training block by image complexity training condition in exposure training condition. ....	67
Figure 7.Total number of exception responses by transfer type across training complexity condition for exposure training group.....	69
Figure 8. Illustration of interaction between exception response pattern and training image complexity.....	70
Figure 9. Illustration of interaction between transfer task and exception response pattern.....	71
Figure 10. Overall performance across training between training conditions. ....	74
Figure 11. Transfer task performance between training conditions. ....	75
Figure 12. Percentage of correct responses in transfer tasks across training image complexity condition. ....	77
Figure 13. Percentage of overgeneralized responses across training blocks by training image complexity condition for the exposure training condition.....	109
Figure 14. Percentage of correct exception responses across training block by image complexity training condition in exposure training condition. ....	109

## LIST OF TABLES

Table 1. Side-by-side comparison of original ATRIUM stimuli and simple stimuli used in the current study.....	26
Table 2. Strengths and weaknesses of single exposure and two image discrimination training perceptual training methods.....	37
Table 3. Matrix of category structure for relevant task dimensions in the perceptual training task. Rows represent stimulus values along the primary dimension, change in glideslope. Columns represent stimulus values along the secondary dimension, PAPI configuration. Training stimulus is designated by a cell containing either an H or an L depending on which side of the primary category boundary the stimulus falls. The two exception categories are labeled EH and EL.....	49
Table 4. Complexity manipulation table. Illustrates the irrelevant stimulus dimensions for each level of complexity. ....	53
Table 5. Population frequency for participant conditions.....	60
Table 6. Means, standard deviations and inter-correlations of primary study variables. ....	60

# CHAPTER 1: INTRODUCTION

## Statement of Problem

The acquisition of visual perceptual skill is no trivial matter. In complex task domains, rapid and precise visual perception can often mean the difference between success and failure, and in some cases, between life and death. On the battlefield, a forward observer's visual perceptions walk the line of successful engagement of the enemy and the disastrous consequences of collateral damage. In medicine, the ability to perceive irregularities on a body scan (e.g., PET, CT, MRI) can lead to early detection and treatment of life-threatening diseases. An airline pilot's accurate perception of the environment out-of-the-cockpit can separate between stable and unstable flight conditions. Moreover, an individual's reliance on visual perception in these domains can be complicated by naturally occurring stimulus distinctions at or near the signal-to-noise threshold (e.g., identifying miniscule abnormalities on a cancer screening X-ray), purposeful attempts to deceive the perceptual system (e.g., camouflaged targets in military domains or hidden contraband in luggage), or even instances where human perception is simply not well adapted to the task environment (e.g., pilot judgments based on perceptions at oblique aerial viewpoints at large distances or making time-to-contact judgments while driving a car at high speeds).

Within these visually complex task domains, successful perceptual performance is predicated on the ability to recognize critical cue features in the task environment while ignoring

the myriad of features that are either irrelevant or even distracting. In stimulus-rich visual environments, task domain experts can perceive critical variations that would otherwise remain imperceptible to the untrained eye (Klein & Hoffman, 1992). This expert capacity to perceive is attributable to the ability to discriminate between fine-grained cue details. Development of these fine-grained perceptual skills is a matter of interest in both the scientific and training development communities. Researchers continue to debate the duration necessary and the machinations underlying the process of perceptual learning in the hopes of unlocking a universal remedy of training. Unfortunately, the complexity of perceptual processes, and by virtue, perceptual learning, has proven a more difficult landscape to navigate as no such panacea has yet been identified.

Beyond suggesting prolonged and repeated exposure to relevant stimuli, there have been few suggestions for how best to train perceptual skill. Most task domains within which perceptual training would best apply involve visually complex environments, rich with both relevant and irrelevant stimuli. Attempts to expose trainees to even a representative sample of possible contexts they may encounter would be a mammoth undertaking. Provided that time and resources are at a premium in most industries, perceptual training is often overlooked or underutilized in favor of more efficiently trained topics like procedural skill or declarative knowledge. With respect to perceptual training, however, one suggested method of improving efficiency is *discrimination training* (Cooper & Podgorny, 1976). By asking individuals to actively discriminate between two relevant stimuli, theorists posit a more robust perceptual



learning experience than exposure alone provides. Discrimination training has been met with promising results in the context of a visual search paradigm (Doane, Alderton, Sohn, & Pellegrino, 1996; Doane, Sohn, & Schreiber, 1999). However, in many complex task domains, the essence of visual search, identifying the presence or absence of a cue in the environment, is simply not sufficient to ensure task success. Instead, perceptual judgments are made based not only on the presence or absence of cues, but also on the degree to which they fit into relevant categories for task completion. Further exploration of discrimination training utility in complex domains can provide insight into the scope within which perceptual discrimination learning is most effective.

### Purpose of Study

Practically speaking, industry training developers have two competing goals, maximization of task learning and minimization of time to accomplish mastery or a predetermined level of proficiency. In complex visual task domains, these needs have rendered traditional perceptual training methods like exposure training ineffective. Discrimination training, however, may be an alternative method of perceptual training that adheres more closely to training developer needs. At the same time, to keep training development firmly grounded in science it is important to be able to base the outcomes of discrimination training on relevant theory. Although there have been instances where discrimination training has been applied in complex task domains (Fiore, Scielzo, Jentsch, & Howard, 2006), to date, there has been no

attempt to validate discrimination training to formal models of category learning in the context of perceptual judgment tasks.

The purpose of this dissertation was to validate discrimination training in a complex task domain with a model of category learning. Provided that most models of category learning have been validated using simple stimuli, an attempt to validate discrimination training in a complex task domain with a model of category learning served two major purposes: (a) to validate a potentially efficient perceptual training technique, i.e., discrimination training, for a more complex perceptual task, and (b) to test and extend a relevant model of perceptual, category learning in a more applied context, i.e., visual approaches in aviation. Pursuing both purposes served to enhance our understanding of perceptual category learning, and in turn, further identify the scope of utility that perceptual training may hold in complex task domains.

## CHAPTER 2: LITERATURE REVIEW

The adage that practice makes perfect echoes throughout nearly every fathomable context of learning science. Successful performers in domains as disparate as athletics (Kalinowski, 1985), music (Gruson, 1988), writing (Scardamalia & Bereiter, 1991), and wine tasting (Melcher & Schooler, 1996) all achieve success through repeated practice of critical skills in their respective disciplines. In fact, there is little, if any, scientific argument against the merits of practice on performance improvement. Through deliberate practice, individuals develop the ability to recognize relevant cues, discriminate fine task distinctions, anticipate future states of the task environment, and generate strategies for task completion. Klein and Hoffman (1992) advocated that development of these perceptual-cognitive skills helps distinguish expert and novice performers. Research, however, suggests that even with deliberate practice, individuals require somewhere in the range of a decade to achieve a true expert level of performance (Chase & Simon, 1973; Ericsson, Krampe, & Tesch-Romer, 1993).

In domains like the ones listed in the above paragraph, expertise is often a selection criterion. That is, entry into these domains is often predicated on demonstration of expert skill. Individuals striving to get into these fields spend considerable time practicing the skills necessary to succeed. As a result, the workforce in these domains is comprised of top performers with substantial prior exposure to relevant tasks. Training professionals in these domains have the benefit of working with individuals who have already contributed a large portion of the time commitment necessary for achieving expertise. Unfortunately, task domains that have less

selective membership requirements cannot rely on extensive prior task experience when considering training. In contexts involving complex task environments (e.g., aviation, military, etc.) this presents a challenge. Due to limited resources for training and the extensive time required for gaining perceptual-cognitive expertise, training in complex task environments often focuses on more practical procedural skill development. Although this training generally prepares trainees for the standard procedure of the task, it leaves the perceptual-cognitive skill development to on-the-job-training or experience. According to those who subscribe to the deliberate practice method for expert development, however, on-the-job experience is simply not sufficient for producing expertise (Ericsson, Krampe, & Tesch-Romer, 1993).

Training perceptual-cognitive skills has long been overlooked, in part, because of the perceived inefficiency associated with it, but also due to a lack of understanding of the underlying nature of perceptual learning. Recent contributions to the perceptual learning literature both in terms of theory and methodology have reinvigorated the discussion on the potential benefits of perceptual training. There still remains a gap between the practical application and theoretical mechanisms suggested to drive perceptual learning however. The research discussed in this document aimed to bridge some of the gap between theory and application of perceptual learning research by couching an applied perceptual training method in a computational model of category learning.

## Perceptual Learning

Historically speaking, the study of perception in psychology can be traced back to the earliest days of the field. From Wilhelm Wundt, to Gestalt theory, to modern day investigations of perceptual differences between real and virtual worlds, perception has maintained a long and continual presence in psychological research. Considering this long-standing history, it is surprising how relatively little is known about how perceptual ability develops. In the early 1960s, James Drever (1960) and Eleanor Gibson (1963) published reviews of perceptual learning that helped ignite discussion of what constitutes perceptual learning. The Gibsonian ecological perspective states that perceptual learning is “any relatively permanent and consistent change in the perception of a stimulus array, following practice or experience with this array” (1963, p. 29). This principally bottom-up perspective of perceptual learning still has many advocates; however, it neglects to acknowledge top-down cognitive processes that impact perceptual learning. In some circles, the mere suggestion that cognitive processes be included in a “true” definition of perceptual learning would cause uproar however. In fact, it is easy to get tangled within the web of semantic definitions that theorists have developed to explain and separate perceptual and cognitive learning. In contrast to this line of reasoning, Goldstone and Barsalou (1998) argued that perception and conception are inextricably linked and therefore should not be explicitly isolated.

Goldstone’s (1998) definition of perceptual learning suggests that enduring changes to the perceptual system facilitate one’s ability to respond to the environment; that perceptual learning helps tailor information gathering processes to the way the information is used. This

definition supports the notion of an interactive symbiosis between perceptual and cognitive mechanisms. Taken further, Hoffman and Fiore (2007) suggest that beyond perception of cues, meaningful integration must take place to elicit perceptual learning.

### Mechanisms of Perceptual Learning

Based on the preceding definition of perceptual learning, the discussion should necessarily turn to perceptual-cognitive processes that underlie perceptual learning. While the logical assertion is that there are multiple processes involved in perceptual learning, the challenge is identifying what these processes are and how each impacts perceptual skill development. Goldstone (1998) identified four mechanisms (i.e., attentional weighting, imprinting, differentiation, and unitization) that cut across the broad landscape of perceptual learning. Although it would be premature to label these as the comprehensive list of mechanisms, each offers rational explanations for separate and sometimes related aspects of the perceptual learning process. Taken together, these four mechanisms help provide a solid theoretical base with which to further explore methods of perceptual training.

### *Attention Weighting*

There is nothing ambiguous about viewing simple one-dimensional stimuli. Making determinations based on perceived differences is straight forward; either the dimension changed or it did not. Unfortunately, in our daily lives, there are few, if any, simple one-dimensional stimuli of which to draw useful information. Instead, everything around us is comprised of combinations of multiple dimensions which provide varying degrees of useful information.

*Attentional weighting*, therefore, refers to the ability of individuals to attend to the relevant information in the environment. In terms of perceptual learning, attention weighting implies that over time, individuals learn to shift attention to relevant cues in the environment (Nosofsky, 1986). Livingston and Andrews (1995) demonstrated this by illustrating the tendency to emphasize cues that reliably predict category distinction during category learning. In contrast to identifying relevant cues, attention weighting is equally impacted by the ability to ignore irrelevant cues in the environment (Haider & Frensch, 1996). Any adjustments to attentional weights can be thought of as changes to a multidimensional similarity space (Livingston, Andrews, & Harnad, 1998). This has profound implications for categorical perception which will be discussed more prominently later in this dissertation.

### *Imprinting*

Past experience is a critical element of perceptual learning in any context. One mechanism through which past experience is manifest in perceptual learning is *imprinting*. Imprinting involves the development of internal detectors for stimuli in the environment (Goldstone, 2003). One of the key points of discussion on imprinting is what precisely is imprinted on the detectors. Using a dot numerosity task, Palmeri (1997) for example, found that individuals developed automaticity on dot patterns which they had been exposed to in training, but not for similar patterns with additional dots. The findings taken further, suggest that there are receptors which imprint whole stimuli. Whole stimuli imprinting supports exemplar based models of category learning (Nosofsky, 1986) which will be expanded upon later in this chapter.

In addition to whole stimuli imprinting, studies have provided evidence that imprinting may also occur on parts of stimuli or patterns within a set of multiple stimuli. Research suggests that in some instances, individuals get better at categorizing stimuli when specific diagnostic features of the stimuli are learned (Schyns & Rodet, 1997). On the other end of the spectrum, there is also research suggesting that imprinting occurs at a more abstract, spatially organized level (Goldstone, 1998). Despite the diverging evidence of these studies, overall, the idea of imprinting implies that there are internal detection mechanisms which are shaped by experience that help optimize processing of repeated stimuli. This line of reasoning closely resembles neurological studies that suggest that a cascading hierarchy of receptors at the cortical level drives perception. Early perceptual processing begins with simple stimulus cell receptors (i.e., line segment) and triggers receptor stimulation progressing to more complex combinations of features (i.e., faces) (Ahissar, Nahum, Nelken, & Hochstein, 2009).

### *Differentiation*

In perceptual learning, the ability to distinguish among task relevant stimuli can be invaluable. The ability to make fine distinctions between stimuli that are seemingly alike to the untrained eye separates expert and novice performers (Hoffman & Fiore, 2007). Differentiation is based on the notion that we learn to perceive as a function of enhanced perceptual precision over time (E. J. Gibson, 1963). Perceptual qualities, features, and dimensions become more distinguishable as perceptual learning occurs. Essentially, the world around us becomes richer with perceptual properties as objects get more distinctive (Gibson & Gibson, 1955). As a perceptual learning mechanism, differentiation can be thought of as an outcome of the previously



mentioned mechanisms. Improved ability to differentiate stimuli is both a function of knowing the critical features to attend in the environment (attentional weighting), and through repeated practice (imprinting) with relevant stimuli.

Like the previous mechanisms discussed, differentiation can be thought of at different levels of perceptual abstraction. Using simple perceptual stimuli, differentiation can be thought of as identifying whether a distinguishing feature is present or absent from a stimulus. In many cases, these simple perceptual differentiation tasks are broken down at the cellular receptor level where differentiation is mostly found to be highly task specific (Fahle & Edelman, 1993). In more practical perceptual applications however, tasks are comprised of much more complex stimuli in which differentiation involves combinations of multiple features, increasing specificity on category and perceptual dimensions (Goldstone, 1998). Further, the spectrum of possible perceptual encounters is not only large, but also varies in occurrence. As a result, the differentiation mechanism, at more complex levels of application, has the dual purpose of reinforcing the separation between commonly occurring stimuli, and developing strategies for correctly distinguishing rare or unexpected stimuli that may be difficult to discern from common stimuli.

### *Unitization*

Up to this point, the focus of the perceptual mechanism discussed in this paper has been primarily on the deconstruction of relevant stimuli into parts. There is little argument that in order to perceive at an expert level, attention weighting, imprinting, and differentiation of fine

grained stimulus characteristics is important. Unitization, however, is a mechanism of perceptual learning which on the surface seems to suggest the development of an opposite skill. Unitization involves the process of constructing a single functional representation of a complex configuration (Goldstone, 1998). Instead of breaking stimuli into parts, unitization suggests that a representative whole is created out of the sum of perceptual parts. Gauthier and Tarr (1997) supported this by examining the effect of prolonged exposure to complex novel stimuli. Their findings suggest that exposure leads to the development of viewpoint-specific representations of stimuli. In line with this, Czerwinski, Lightfoot, and Shiffrin (1992) described the unitization process in terms of chunking visual features together. Developing unitized representations is especially helpful for task stimuli which have commonly occurring features that, when combined, consistently require the same response.

### *Summary*

To gain a better understanding of perceptual learning, it is important to consider the underlying mechanisms which drive perceptual skill development. The four mechanisms discussed do not necessarily constitute a comprehensive list of mechanisms, but provide an informed point of reference to guide the discussion toward training theory and methodology. Among these mechanisms, there are obviously overlaps which make it difficult to proclaim one more impactful than others. Instead, the implication is that the task itself serves as the conduit through which appropriate mechanisms are employed in a perceptual task. Overall, the mechanisms described can be thought of as residing on a continuum of feature specificity. This is illustrated in the general propensity for these mechanisms to support either the breakdown of

stimuli into parts or integration of features into a whole to demarcate perceptual skill development. The implications of the distinction of mechanisms again make it difficult to generate a panacea recommendation for training perceptual skill, but provide an informative look at mechanisms to consider when developing training.

### Perceptual Judgment

Since the discussion of perceptual learning inhabits the often debated gray area between perception and cognition, there is no definitive evidence for where perceptual learning begins or ends. As a result, perceptual skill can be manifest in a number of different relevant perceptual tasks depending on how it is interpreted. This is evident in the distinction between visual search and perceptual judgment tasks. Visual search is characterized as a process in which decision outcomes are based on identifying the presence or absence of target stimuli. There is little argument that perceptual processes influence the ability to perform visual search. Perceptual judgments, on the other hand, also occur when presence or absence is not the determining factor of the task. Perceptual judgment can be thought of as a judgment of the magnitude of states (e.g., size, distance, weight, orientation) of perceptual stimuli that are present in the environment (Tajfel, 1957). Perceptual judgment is less contingent on whether something *is* perceived and more contingent on *how* it is perceived. As a result, the argument can be made that perceptual judgment is more cognitive than perceptual in nature. In fact, perceptual judgment is often equated with the ability to categorize stimuli along relevant dimensions. That is, perceptual judgment skill is not necessarily contingent on deriving precise estimates of a stimuli state, but rather hinges on being able to accurately assign stimuli to categories which inform task decision

making. For instance, a pilot does not need to calculate distance/altitude measurements to within a few meters of accuracy when making a perceptual judgment during a visual approach to land. Instead, a pilot should be able to recognize states of being too high or too low for stable approach accurately enough to maintain a safe descent path. Although the more liberally defined concept of perceptual learning which this dissertation uses could certainly include category learning, categorization has largely emerged as a separate field from perception (Op de Beeck, Wagemans, & Vogels, 2003). Provided that there are similarities in the perceptual processes that contribute to both visual search and perceptual judgment, it is logical that some of the training requirements to achieve skill in both may also be shared. The following section of this chapter further explores category learning. Offering evidence for the linkages between perceptual learning and category learning strengthens the assertion that perceptual training has utility in perceptual judgment tasks.

### Category Learning

Humans have the propensity to sort the vast array of information that surrounds them. The process of sorting information can yield a wide range of physical (e.g., large, round), functional (e.g., sharp, slow), and intangible (e.g., angry, good) labels of the world. This process of categorization can be thought of as a way to organize information into more manageable pieces of information. Although categorization and perception are often discussed as separate processes, research supports the idea that these processes interact (Op de Beeck, Wagemans, & Vogels, 2003). In relation to perception, there is evidence that categorization leads to sensitized response to category-relevant dimensions, de-emphasis of category irrelevant variations, and

selective sensitization of relevant dimensions at category boundaries (Goldstone, 2003). Due to the influence of categories on perceptual processes, it is reasonable to surmise that category learning can lead to improved performance of perceptual judgment tasks.

In general, the process of category learning is most closely linked to perceptual differentiation. Categories are formed by perceived differences of dimensions of stimuli. Category learning can take place in novel tasks where the formation of new categories occur, or in more familiar tasks where categories are tuned to fine details that are imperceptible to less trained individuals. Each category can be comprised of separable dimensions (e.g., size and color), where dimension variation can be easily separated from others, and integral dimensions (e.g., saturations and brightness) where variation between dimensions are fused (Op de Beeck, Wagemans, & Vogels, 2003). These dimension characteristics often influence the complexity of categorization. Whereas stimuli comprised of separable dimensions can be attended selectively, integral dimensions are less easily attended in isolation. This may suggest that stimuli with highly separable dimensions are processed into categories by breaking dimensional features into parts (i.e., selective attention, differentiation). Along the same line of thinking, stimuli with highly integral dimensions may be processed into categories by combining dimensions into more holistic representations (i.e., unitization). These distinctions of category learning (i.e., task familiarity and dimension partition) are important considerations when determining the nature of category learning for specific tasks.

The general conception of category learning is relatively straight-forward. Similar to perceptual learning, however, the underlying mechanisms that drive category learning are less easily identifiable. This has resulted in the development of numerous models that attempt to explain category learning. In the following section the discussion will turn to models of category learning. By first briefly describing the types of models that have been conceived, the discussion will naturally lead into a model for testing methods of perceptual training.

### Models of Category Learning

There are two general types of models which are prevalent in psychological research, conceptual models and computational models. Conceptual models are representations of how topic-relevant concepts interrelate to form a psychological process. In complex multi-layered psychological processes, like category learning, where the concepts that make up the process are numerous and often a matter of debate, it is difficult to derive a comprehensive representation of the whole process. Due to the complicated nature of category learning, many researchers have turned to computational modeling to help explain behaviors associated with the construct (Kruschke, 2008). Computational models are helpful for modeling behavior where the number of relevant concepts is not as easily defined. In computational models, simulations of performance are generated to compare to real world data. This is accomplished through the development of, at times, complex mathematical equations which serve as a more abstract representation of mechanisms underlying the construct in question. Using this modeling technique, a number of different theories of category learning have emerged, including exemplar, rule-based, prototype, and boundary models (Kruschke, 2008). While each of these theoretical bases have produced

informative lines of research that have helped to advance understanding of category learning, two in particular, exemplar and rule-based models of category learning, are particularly well suited for explaining the perceptual learning mechanisms that occur in perceptual judgment tasks. As a result, these two types of models will be the focus of discussion. Both exemplar and rule-based theories of category learning can be considered single system theories.

### *Exemplar Models of Category Learning*

In the late 1970s, Medin and Schaffer (1978) introduced a theory of category learning known as the context theory. The context theory served as a slight departure, at the time, from more prevalent prototype theories of categorization by suggesting that category judgments are made based on the similarity of stimuli to exemplars stored in memory. This theory served as the catalyst for the development of modern exemplar models of category learning such as the General Context Model (GCM; Nosofsky, 1986), Supervised and Unsupervised Stratified Adaptive Incremental Network (SUSTAIN; Love, Medin, & Gureckis, 2004), and the Attention Learning Covering Map (ALCOVE; Kruschke, 1992). Similar to the concept of whole stimulus imprinting, exemplar models presume that all stimuli that an individual has been exposed to are stored as exemplars with category labels in multidimensional memory space (Ashby & Maddox, 2005). One argument against exemplar theories is that any one exemplar may fall into several different categories that are contingent on a different combination of features that make up the exemplar. In order to distinguish the appropriate category, attending to and ignoring features, which according to the theory, are not stored individually, is necessary. To address this, GCM includes a selective attention mechanism which allows for individuals to

direct focus to category relevant features within exemplars (Nosofsky, 1986). The GCM has been found to fit data in a wide variety of category learning contexts, but still lacks any type of mechanism for learning that occurs with repeated exemplar exposure. The ALCOVE model (Kruschke, 1992), which is a direct descendant of the GCM provides this learning mechanism.

Similar to the concept of imprinting, exemplar theory is built on the notion that as exposure to stimuli increases, so does the amount of category learning. Exemplar models predict that optimal performance will eventually occur with exposure (Ashby & Maddox, 2005). In terms of training, this would most closely resemble the benefits of exposure training. Most critics of exemplar theory as a single system theory of category learning point to the fact that it does not do an effective job of accounting for the extrapolative nature of category learning. In exemplar theory, after each stimuli exposure, a memory store is created. The more times that exemplar is accessed (the stimulus is perceived), the more quickly and accurately that the exemplar can be categorized. This does not account for instances where novel stimuli are encountered. As a result, exemplar theory most closely resembles interpolated skill which is described in more detail later in this chapter.

#### *Rule-Based Models of Category Learning*

Rule-based models of category learning offer a slightly different perspective of category learning. Unlike exemplar theory where there are representation stores of every exposed stimulus, classic rule-based models of category learning instead are based on the premise that every category (not stimulus) is stored as a list of necessary and sufficient features (Smith &



Medin, 1981). These category feature lists are used as a comparison guide of stimulus features. Stimuli are categorized by how well their features match with task relevant category lists. That is to say that any stimulus may fall into a number of categories depending on the context in which it is being referenced.

Rule-based models offer an alternate viewpoint of category learning in which extrapolative skill, which is described in more detail later in the chapter, can be explained. Rules are not tethered to specific exemplar stores, but instead, are made up of relevant lists of features. In this respect, individuals can apply rules to stimuli which they have not previously been exposed. The feature list comparisons that make up rules are similar to the differentiation mechanisms described for perceptual learning above.

As a stand-alone theory of category learning, the classic rule-based model holds up well in instances where conjunctive (i.e., “AND”) rules are sufficient for categorization (e.g., a stimulus is in category A because it is tall *and* wide; Ashby & Maddox, 2005). Unfortunately, in many complex domains, categories are not as simply explained as is the case with conjunctive rules. Disjunctive (i.e., “OR”) rules are based on the notion that categories could be made up of stimuli in which there is no uniform combination of features that predict category membership (e.g., a stimulus is in category A because it is tall and wide, *or* because it is short and narrow). In disjunctive rule categories, the presence or absence of a single feature cannot be used to guide category prediction. This makes the conception of the list of necessary and sufficient features more difficult to explain. Newer conceptions of rule-based category learning have incorporated

an exception learning mechanism to offset potential disjunctive category labels (e.g., RULEX; Nosofsky, Palmeri, & McKinley, 1994). In spite of this, the predominant modern viewpoint of rule generation as a single system explanation of category learning have been supplanted by the concept of category learning involving multiple systems.

### *The ATRIUM Model*

Although the range of single system models offers serviceable explanations of category learning while maintaining a certain degree of parsimony, none are without limitations to the categorization behavior they explain. In many cases, the shortcomings of one model are accounted for in another model. This suggests that perhaps the quest for parsimony is clouding the complexity of category learning. Instead of offering a single system explanation, more recently theorists have turned to combining models to explain category learning (Ashby & Maddox, 2005). Although these multi-system models are more complex, and require more complex computations, they are better able to explain disparate aspects of category learning. Exemplar and rule-based models alone are unable to provide a complete picture of category learning, in a multi-system model however, they are complementary. The strengths of one offer an explanation for the weaknesses of the other. This complementary relationship in combination with a semblance to the perceptual skills discussed earlier, suggest that a multi-system model of category learning may help to bridge the scientific gap between perceptual training methodology and learning associated with perceptual judgments. One such model is the Attention To Rules and Instances in a Unified Model (ATRIUM) as proposed by Erickson and Kruschke (1998).

The ATRIUM model is a computational model which combines the thoroughly tested ALCOVE exemplar model with a rule-based module. The model is driven by five psychological principles: exemplar representation, rule representation, representational attention, dimensional attention, and error-driven learning (Erickson & Kruschke, 2001). These principles are housed in four components which comprise the ATRIUM model: a rule module, an exemplar module, a gating mechanism, and learning. In the following sections, a brief description of the components of the ATRIUM model will be explained in terms of these psychological principles.

#### *Rule module*

The rule module makes use of the rule representation and dimensional attention principles on which ATRIUM is based (Erickson & Kruschke, 2001). Under the rule module, stimuli are classified according to adherence to dimension specific rules in which stimuli fall. This can be thought of in the most basic conception in terms of boundaries. Each category is separated by a boundary that is determined by a combination of relevant dimensions; the dimensions, in turn, dictate category membership. Following this path, each dimension which impacts category membership can be separated by boundaries driven by the development of rules relevant to the dimension (e.g., long, short). Logically, as a stimulus gets closer to rule boundaries, the more difficult it may be for a rule to predict category membership and likewise up the hierarchy (near boundary categorizations are more difficult than far boundary categorizations).

The ATRIUM rule module is actually made up of multiple individual modules that are each specific to only one dimension of the stimuli. This dimension separation across modules allows for the rule module to make use of the principle of dimensional attention, by isolating and weighting category critical dimensions. Dimensional attention suggests that people learn to attend to relevant dimensions on the stimuli. That is to say that multi-dimensional stimulus may have rule nodes that are more predictive of category membership depending on the context. With multi-dimensional stimuli, different combinations of rules may yield different categorizations. As a result ATRIUM is set to learn multiple boundaries on a single dimension, depending on other dimension inputs.

#### *Exemplar Module*

The exemplar module, as mentioned above, is a complete replication of Kruschke's ALCOVE model (1992). This module employs both the exemplar representation and the dimensional attention principles listed above (Erickson & Kruschke, 2001). Under the exemplar module, stimuli are classified according to similarity to stored exemplars which are members of relevant categories. Conceptually, this can be thought of in terms of nodes. There are a series of exemplar or category nodes which are activated by a set of dimensional nodes that are dictated by currently perceived stimuli. Similar to the rule module, dimension nodes serve to guide dimensional attention. Where the module differs in this respect, is that dimensional nodes are not separately activated by individual rules. Instead, each relevant exemplar node, which can be thought of as full representations of previously exposed stimuli, integrates all dimensional node information. Also, instead of the dimensions being parsed by rules, the exemplar module learns

which dimensions are more pertinent to the categorization task. This helps to focus dimensional attention on the most important dimensions of the overall stimuli for categorization. This dimensional stimulus input helps dictate which exemplar nodes are most similar to the stimuli. Based on the activated exemplar nodes, category nodes are, in turn, activated using the same principle of similarity.

### *Gating Mechanism*

One of the challenges of combining multiple systems into one model is in determining the contribution of each system to the task. In ATRIUM, this is accomplished by linking the two modules with a mathematical gating mechanism (Erickson & Kruschke, 1998). The gating method capitalizes on the principle of representational attention. That is, people learn to use different representations (i.e., exemplar and rule) depending on how well suited each is to the particular stimuli (Erickson & Kruschke, 2001). One way of looking at this is that the previous modules serve to classify stimuli, whereas the gating mechanism serves to classify appropriate modules for optimal categorization.

The gating mechanism in ATRIUM is based off of Jacobs, Jordan, Nowlan, and Hinton's (1991) algorithm. Using this method, every stimulus used is processed by the rule module and exemplar module simultaneously. The contribution of each module is dependent upon the activation strength within each module. Activation depends on a combination of the stimulus proximity to rules or exemplars in space in addition to the level of learning that has occurred

(Erickson & Kruschke, 1998). The results of the gating mechanism result in a set of cost parameters that indicate the benefit of each module for classification of the particular stimuli.

### *Learning*

No model of category learning is complete without the inclusion of the concept of learning itself. The final psychological principle at work in the ATRIUM model is the concept of error-driven learning. Error-driven learning is based on the idea that individuals learn from feedback received on incorrect responses to items (Erickson & Kruschke, 1998; Erickson & Kruschke, 2001). In the ATRIUM model, each module (including the gating mechanism) learns by incrementally adjusting behavior so that repeated presentation will increase the likelihood of a correct response. Learning is determined by the amount of feedback received in each module. Successful classification within one of the modules impacts the level of feedback that will be received in future instances. If a module performs more poorly, the feedback will decrease as the overall model attempts to optimize classification strategy.

### *Summary of ATRIUM*

The ATRIUM model is comprised of a complex set of mathematical equations which offer a quantitative explanation for category learning. Despite the quantitative nature of ATRIUM, the conceptual underpinnings hold a number of parallels to some of the mechanisms discussed in regard to perceptual learning. The idea of complete stimulus storage that is strengthened by repeated exposure in the exemplar module closely resembles the mechanisms of imprinting and unitization as presented by Goldstone (1998). Likewise, the concept of

differentiation which suggests that perceptual learning involves breaking stimuli down into differential parts in order to identify them, closely adheres to the rule module which processes each dimension of a stimuli separately for categorization. Additionally, both the principles of representational and dimensional attention resemble the perceptual learning mechanism of attentional weighting. Provided that learning is the central tenant which drives both perceptual and category learning models, it is easy to see where there is great conceptual overlap. As a result, the ATRIUM model serves as a suitable theoretical framework from which attempts to test the effectiveness of a specific perceptual training technique.

#### *Effect of Stimulus Complexity on Model Fit*

To date, models of category learning have been tested using only simple stimuli and only using active exposure training method. While informative, applying the results of this research into the complexity of real world perceptual and category learning tasks is not clear cut. Although simplifying stimuli to more easily isolate behavior is a necessary step, it is rare that complex domains offer stimuli with as readily discernable dimensions. In addition to each relevant dimension of stimuli in complex environments, there are countless irrelevant cues that are not included in basic tests of category model fit.

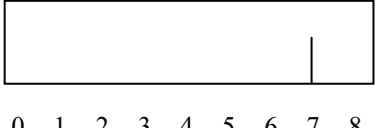

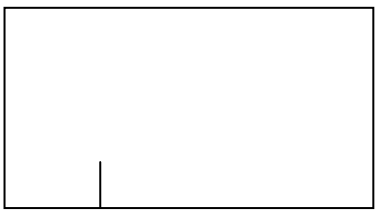

The current study therefore had two purposes: First, study how well the theory (i.e., the ATRIUM model) holds up in the face of increased stimulus complexity. Second, investigate discrimination training and compare it to the model. With respect to the former, I developed, an applied setting in which a target categorization task can closely resemble the previously used

simple categorization task (see Table 1). The visual approach task, in its simplest form, fulfilled this need. I then extended the stimulus environment to more realistic settings by step-by-step adding additional stimulus classes found in the real-world task.

Table 1. Side-by-side comparison of original ATRIUM stimuli and simple stimuli used in the current study.

Original ATRIUM Task		Visual Approach Task	
Primary Cue (PC): Rectangle Height		Primary Cue (PC): Position of Runway Line	
Secondary Cue (SC): Line Position		Secondary Cue (SC): Color Combination of Dots	
Original Rule	Original Atrium Stimuli	Applied Task Rule	Proposed Simple Task Stimuli
If PC > 4.5 then <b>A</b>	<p>A 9x9 grid with a vertical line at x=4.5 and a rectangle from y=1 to y=7.</p>	If PC > 3° then <b>A</b>	<p>A vertical line at x=7.5 and three dots at x=7.5, y=1, 2, 3.</p>
If PC < 4.5 then <b>B</b>	<p>A 9x9 grid with a vertical line at x=1 and a rectangle from y=1 to y=4.</p>	If PC < 3° then <b>B</b>	<p>A vertical line at x=1 and four dots at x=1, y=1, 2, 3, 4.</p>
<b>Original Rule Exception</b>		<b>Applied Task Rule Exception</b>	



<p>If PC + SC = 9 when PC = 2 and SC = 7 then <b>EA</b></p>		<p>If PC = 4.5° and SC = WWBB then <b>EA</b></p>	
<p>If PC + SC = 9 when PC = 7 and SC = 2 then <b>EB</b></p>		<p>If PC = 1.5° and SC = BBRR then <b>EB</b></p>	

I expected that the data resulting from a test with the minimal stimulus equivalent of the visual approach task would produce similar response patterns to the original ATRIUM task, not just with respect to the percentage of correct responses, but also as a proportion of rule and exception responses. Including overall performance and rule/exception responses provides a way of studying the applicability of the ATRIUM model with respect to rule-based and exemplar-based response strategies. To get the most representative picture of model fit, it is important to look to training performance to confirm that the appropriate category structures were learned, as well as transfer performance measures to gauge the extent that the skill applies to relevant tasks.

As discussed earlier, the ATRIUM model is based on the idea of both exemplar and rule-based mechanisms impacting category learning. According to the model, the contribution of each mechanism is dictated by the activation strength produced by each stimulus. The result is a trade-off system where when one mechanism has greater activation, it is more likely to be utilized to make category distinctions. In simple stimulus category learning tasks, early category learning is characterized by what is known as overgeneralization. Overgeneralization is the propensity to more frequently label items based on rules even in instances of rule exceptions. Over time, exemplar based processes become more influential as exemplar associations are strengthened. The result is an appropriate balance of rule based and exemplar based influence on category judgment. As such, it is logical to deduce that adding stimulus dimensions, regardless of their relevance to the task, will shift the response pattern. The process can be thought of in terms of attentional weighting, in which over the course of training irrelevant information will be learned to be suppressed. The resulting suggestion would be that the influence of rule-based mechanisms will be more pronounced for longer in training. As such the following hypotheses reflect the predicted shift based on level of training stimulus complexity.

*Hypothesis 1: As training stimulus complexity increases, ATRIUM model fit will hold, such that participants will rely more heavily on rule-based response strategies.*

*Hypothesis 1a: As training stimulus complexity increases, participant reliance on rule-based response strategies will increase such that overgeneralization will be evident in training for more trials than in simple stimulus conditions.*

*Hypothesis 1b: As training stimulus complexity increases, participant reliance on rule-based response strategies will increase such that a lower proportion of exception training stimuli will be labeled correctly.*

*Hypothesis 1c: As training stimulus complexity increases, participant reliance on rule-based response strategies will increase such that less overall exception responses will be given in the transfer task.*

*Hypothesis 1d: As complexity increases greater reliance on rule-based processes will impact dimensional attention in the transfer task such that a higher proportion of exception responses will occur on the primary (rule) dimension of exceptions than on the secondary (exception) dimension.*

### Perceptual Training

One of the ever-present challenges that exist in the research community is translating research findings into practical solutions in relevant work domains. Merely identifying underlying mechanisms of perceptual learning does not guarantee an effortless transition to practical implementation. Stating that one needs practice to develop perceptual mechanism skills is in many respects too general a distinction for an individual without intimate knowledge of training science. Before getting into a specific perceptual training method description, the conversation will turn to perceptual skill development. Providing insight into the process of perceptual and category learning from the previous sections in combination with a discussion of

the resulting skills that accompany these processes will help to direct conversation to perceptual training methodology.

### *Perceptual Skill*

The goal of any training program is to produce outcome skills that improve task performance. Perceptual learning can be thought of as the process by which an outcome, perceptual skill, is shaped. Perceptual skill can manifest itself in a wide variety of resulting domain-specific task improvements. That is, perceptual skill development can improve the ability to spot a card-counter in a casino, hit a baseball, or land an aircraft. Despite the varying nature of perceptual skill across domains, these specific skills can be more generally categorized into two types of skill, interpolated and extrapolated skill. Posner and Keele (1968) were early proponents of this skill distinction, suggesting that the amount of variation in training stimuli will impact whether or not specific or generalized skill is developed.

### *Interpolated Skill*

Repeated exposure to task relevant stimuli has long been recognized as a central tenet of perceptual learning. The result of repeated exposure to a consistent collection of stimuli leads to the development of interpolated skill. Interpolated skill also known as stimulus-specific skill, is closely associated with the concept of automaticity (Shiffrin & Schneider, 1977). That is, repeated exposure leads to strengthened association between stimuli and subsequent responses (Schneider & Detweiler, 1988). After enough exposure, individuals become able to quickly identify and extract meaningful information from stimuli to which they have been exposed

(Karni & Sagi, 1991). This is a particularly helpful skill in domains where there are consistently matched stimulus and response pairs (Logan, 1988). Because interpolated skill is very specific to the stimuli in which exposure has occurred, it is not particularly pliable to previously unexposed material. As task stimuli get more varied and complex, it is logical to deduce that it will take an extended amount of time to develop interpolated skill for all relevant stimuli.

### *Extrapolated Skill*

Whereas interpolated skill is specific to previously exposed perceptual stimuli, extrapolated skill is associated with how well an individual can apply what was learned to novel stimuli or different tasks. Instead of strengthening memory of specific instances, extrapolated skill involves developing strategies that generalize outside of just the exposed stimuli. Research provides evidence that rule based skills are impactful in simple problem solving tasks (Anderson, Fincham, & Douglass, 1997; Haider & Frensch, 1996). These have also been extended to simple perceptual learning tasks as well (Doane, Alderton, Sohn, & Pellegrino, 1996; Doane, Sohn, & Schreiber, 1999). Variation, as opposed to repetition, of training stimuli has been found to elicit development of extrapolated skill (Kerr & Booth, 1978). However, unlike interpolated skill, the manifestation of extrapolated skill may not occur immediately following training (Schmidt & Bjork, 1992). In spite of the initial short delay manifestation of extrapolated skill, in complex stimulus domains, it is reasonable to suggest extrapolated skill development will require less time intensive training to achieve proficiency.

### *Summary*

There is plenty of evidence that suggests that both interpolated and extrapolated skills are critical outcomes of perceptual learning. Using simple two-dimensional stimuli, Erickson and Kruschke (1998; 2001; 2002) have been able to better account for both interpolative and extrapolative skill development in categorization tasks, than single system models. Interpolated skill is specific to previously exposed stimuli, while extrapolated skill is applied more generally to both previously exposed and novel stimuli. The challenge of the training developer is determining what mix of skill is both practical and effective in training the perceptual skill desired. In domains with very predictable stimulus and response sets, interpolated skill can be quite useful. On the other hand, tasks involving ambiguously defined stimulus and response sets will be more suitably prepared with extrapolated skill. In all reality, most task domains do not fall neatly into one or the other category. Using aviation as an example, a pilot may experience countless visually normal approaches to a runway. The repeated exposure in practice itself likely serves to strengthen interpolated skill for identifying these normal conditions. In rare instances visual conditions can generate illusions that may cause abnormal conditions to appear normal. Due to the infrequency of occurrence, extrapolated skill will be more helpful to identify the illusion. Experts in these domains, should not only be able to react quickly to scenarios to which they have previously been exposed, but also effectively shift from interpolated to extrapolated skill response when unfamiliar events occur.

#### *Perceptual Training Method*

Perceptual training has been overlooked in many domains in which it could prove beneficial. This is probably attributable to a number of things, but the most prominent is

efficiency. Even in domains where perception plays an important role, the development of procedural skill yields more training gains more quickly, strictly speaking, in terms of cost benefit. Despite this, perceptual skill remains critical in a large number of task domains. Having provided insight into the process (perceptual learning) and outcome (perceptual skill) the discussion now turns to training method.

The core principle of perceptual training is exposure. In the training context, exposure is defined as presentation of task related stimuli with the intention of strengthening perceptual skill. Although in a controlled setting, the objectives of exposure are clear, exposure can also result from less structured on-the-job experiences as well. Whether recognizing auditory tones (Oakes, 1955), identifying olfactory characteristics of wine (Wilson & Stevenson, 2003), or seeing small feature differences in difficult visual tasks (Biederman & Shiffrar, 1987), training these perceptual skills involves some level of exposing individuals to task critical stimuli. The concept of exposure follows closely with the notion of practice or experience-based learning described earlier in the paper. In fields where there are limited numbers of stimuli which will be encountered, repeated exposure can be helpful for imprinting relevant stimuli. Ultimately, this can lead to individuals developing what seems to be an almost automatic, unconscious reaction to stimuli presented (Shiffrin & Schneider, 1977). As mentioned previously, the practicality of exposure gets called into question as the complexity of the task increases. Although exposure training can conceivably elicit any one of the processes of perceptual learning, it is most closely linked to imprinting. In order for exposure to elicit learning gains associated with attentional

weighting or differentiation, individuals have to rely on previously stored information. The additional cognitive task of information retrieval logically would take more time. In addition, training to imprint every conceivable stimulus becomes an issue of time. In order to present perceptual training as a viable training option in complex task environments, training developers must look for ways to augment the benefits of exposure with other training techniques. One approach to this is to apply general techniques (i.e., feedback, difficulty manipulation, etc...) that have been found to enhance learning in other training contexts. Although utilizing these general training techniques can help to fine tune training, for perceptual training, it is perhaps more beneficial to look at ways to more directly tap into the mechanisms which underlie perceptual learning.

### *Discrimination Training*

One method that has been suggested to augment exposure is discrimination training. Discrimination training is a more explicit form of eliciting the process of differentiation in perceptual learning. Other exposure-based training methods generally involve presentation of single stimuli, followed by individuals relying on previous memory of similar stimuli to determine the appropriate responses. The process of differentiation must occur with previously stored representations of similar stimuli. Discrimination training reduces the memory load requirements. Individuals are presented with two stimuli either simultaneously or in succession and asked to determine whether the critical stimuli are the same or different based on task relevant cues. Although the majority of research conducted using discrimination training has involved the use of simple stimuli (Doane, Alderton, Sohn, & Pellegrino, 1996; Fahle &



Edelman, 1993), the findings hold promise for the utilization in a more complex perceptual training environment. Researchers have been able to support the use of discrimination training for stimulus-specific (interpolated) skill (Karni & Sagi, 1991). In particular, the research by Karni and Sagi suggests that discrimination learning is impactful at the early cortical stages of visual processing. From the perspective of exposure training, the penchant for discrimination training to use two images instead of one image increases the rate at which individuals are exposed to task relevant stimuli. It is logical that this alone, may provide additional training benefit of single exposure training for the development of interpolated skill. Taken further, discrimination training has also been found to impact extrapolated skill as well. Using a series of basic shape comparisons, Doane and colleagues (Doane, Alderton, Sohn, & Pellegrino, 1996; Doane, Sohn, & Schreiber, 1999) found that discrimination training has implications for strategy development in addition to stimulus recognition. If individuals are provided feedback on each trial, they can adjust their response strategy based on inferences about how characteristics of the stimuli interact. This suggests that discrimination training not only provides additional exposure to images, but in addition provides direct experience differentiating critical perceptual features for the task. Through this, individuals are also able to make valuable cue comparisons which can lead to strengthened weighting of attention to relevant cues and improved ability to ignore irrelevant cues to the task.

The model which the Doane et al. (1996; 1999) discrimination training research was based on is Fisher's optimal feature model (Fisher & Young, 1987; Fisher & Tanner, 1992). This

is a model of visual search. Visual search, as described above, involves primarily determining the presence or absence of critical stimuli or cues in the environment. Building off of the work using simple stimuli, the discrimination training method has been found to enhance learning in more applied settings like security baggage screening (Fiore, Scielzo, Jentsch, & Howard, 2006). This task environment is still largely defined by visual search for objects in the environment.

In domains like aviation, the perceptual task may be more driven by making a judgment about the current state of the aircraft in relation to the information present in the environment. In complex task environments such as this, simply identifying if the cue is present or absent may not be sufficient for task completion. Instead, it might be more logical to think of the task as a categorization of the situation (e.g., too high, too low, etc...) which guides subsequent behavior. To insure that discrimination training provides the same learning gains in a perceptual judgment task as in visual search, the logical step is to apply discrimination training into a field that is characterized more by perceptual judgment.

#### *Effect of Training Method on Category Learning*

Training tasks associated with category learning are often a single exposure training task. As discussed above, exposure training provides a means for eliciting perceptual learning. In environments rife with relevant and irrelevant stimuli, limitations like the considerable time to reach an acceptable performance range as the number and complexity of stimuli increases, limits the training application of the finding (Table 2). Bearing this in mind, the second thrust of the

proposed study focuses on how training method influences the effectiveness of a perceptual training task in a transfer perceptual judgment task.

Table 2. Strengths and weaknesses of single exposure and two image discrimination training perceptual training methods.

	Single Exposure	Two Image Discrimination
Strengths	<ul style="list-style-type: none"> <li>- Resemblance to real world task</li> <li>- Single exposure facilitates imprinting of images</li> <li>- Repetition of images supports development of exemplar stores (interpolation)</li> </ul>	<ul style="list-style-type: none"> <li>- Alternating between same and different image pairs facilitates attentional weighting to task relevant cues</li> <li>- Same stimulus pairs facilitate development of imprinting relevant cues</li> <li>- Different stimulus pairs facilitate differentiation of relevant cues</li> <li>- Variation of image pairs supports development of rules (extrapolation)</li> <li>- Greater number of images exposed in same training time</li> </ul>
Weaknesses	<ul style="list-style-type: none"> <li>- Greater amount of memory reference for differentiation</li> <li>- Greater amount of memory reference for attentional weighting</li> <li>- Less images exposed in same training time</li> <li>- Less variation between images in same training time limits development of rule development (extrapolation)</li> </ul>	<ul style="list-style-type: none"> <li>- Less focal scan time per individual image slows development of exemplar stores (interpolation)</li> <li>- Changes nature of task (same/different vs. strict categorization)</li> </ul>

Practical relevance in the context of this study is best conceived as how well training adheres to training goals in an applied setting. In perceptual training especially, time to skill proficiency is of interest due to prior work suggesting that perceptual skill requires prolonged periods of exposure (Chase & Simon, 1973). In addition to, and in some respects more obviously, the overall performance gains of the training are also of interest in terms of practicality of training. This is even more important if the training is not a strict replication of the task environment. In these instances, it is important to address how the skill obtained in training

transfers to overall performance gains. Ultimately, the findings of practical relevance are of interest to practitioners interested in applying the research findings into the development of training methods.

As emphasized earlier, the method in which perceptual stimuli are presented can have an impact on the development of interpolative and extrapolative skills. Although the research in this area has been primarily limited to studies involving visual search (Doane, Alderton, Sohn, & Pellegrino, 1996; Doane, Sohn, & Schreiber, 1999), the similarities between visual search-based perceptual learning and category learning suggest that these findings may translate into a perceptual judgment task in a complex environment. The single image exposure method of perceptual training which is commonly used in both visual search and category learning research parallels the concepts of imprinting as a perceptual learning mechanism and exemplar-based learning in a category learning sense. If these parallels hold up in an examination of training methods, one would expect to find that repeated single exposure to stimuli would result in noticeable performance improvements with images that participants have been previously exposed to. Performance would, however, suffer when transfer to previously unexposed items is introduced. A two-image discrimination method of perceptual training may serve to alleviate these extrapolation skill decrements. In addition, the increase in stimulus exposure per item (two instead of one) will lead to a reduced number of items required to achieve the same level of interpolated skill, thus reducing overall necessary training time requirements. By generating an active differentiation task (two images side-by-side), the exemplar/imprinting skill will still be

derived. In addition, though, rule development and testing which may help to extrapolate to previously unexposed stimuli should also take place. This will lead to an improvement in both interpolated and extrapolated stimuli in a transfer task. As such, the following hypotheses are put forth for the effect of training method on category learning.

*Hypothesis 2: Two-image discrimination training elicits explicit practice with both exemplar and rule-based categorization strategies such that it will result in better training outcomes than single exposure in the perceptual training task.*

*Hypothesis 2a: By increasing the number of images per exposure, the discrimination training will result in reduced time to reach training proficiency.*

*Hypothesis 2b: Two image discrimination training elicits explicit practice with both exemplar and rule-based categorization strategies such that overall accuracy will be better on the transfer task than in the single exposure training condition.*

#### *Moderating Effect of Training Method on Training Stimulus Complexity*

As evidenced by the previous hypotheses, both of the manipulations (complexity and training type) previously discussed should have important individual impact on training outcomes and model fit. Provided the previous assertions about the added benefit of discrimination training for training outcomes hold, it is reasonable to assert that this type of training will ameliorate perceptual training in any circumstance. When considering training complexity though, the observed effects of training method may be more pronounced as training

complexity increases. Considering that there are fewer varying visible cues in simple stimuli, there will be less information to suppress and subsequently less irrelevant variation between images. Following this line of reasoning, it is logical to think that the exemplar based associations will strengthen more quickly in the absence of additional information to suppress. In this case the benefits of discrimination training will be more noticeable when complexity increases. The following hypotheses reflect this logic.

*Hypothesis 3: Training method will have a moderating effect on perceptual training such that discrimination training will have a greater improvement of training effectiveness as training stimulus complexity increases.*

*Hypothesis 3a: Two image discrimination training elicits explicit practice with both exemplar and rule-based categorization strategies such that the increased overgeneralization caused by increasing training stimulus complexity will be ameliorated.*

*Hypothesis 3b: Two image discrimination training will produce better overall transfer task performance than single exposure training as complexity increases.*

## Chapter Summary

In this chapter, a theoretical base was provided for perceptual learning, category learning, and perceptual training. By drawing parallels between perceptual learning theory and the more cognitively driven category learning theory, the case for applying perceptual training techniques

into a perceptual judgment task seems logical. Investigating the impact of task complexity, will provide an important link between theory grounded in simple task learning and application into more complex real world tasks. In addition the investigation of training type will help further research involving techniques which may increase efficiency of the task. By investigating these factors in combination (Figure 1) an under-examined piece of the perceptual learning puzzle will be addressed.

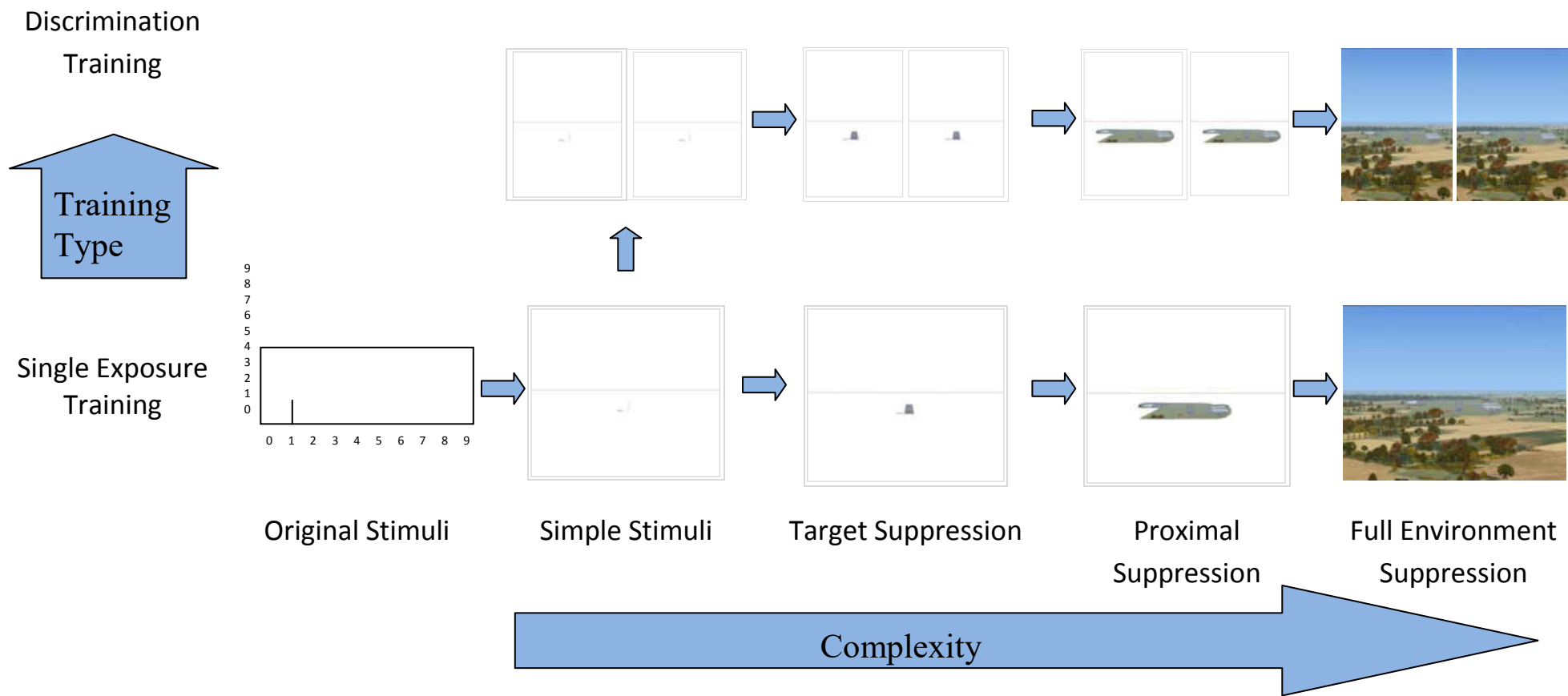


Figure 1. Graphical representation of experimental comparisons.



## **CHAPTER 3: MATERIALS AND METHODS**

### **Participants**

A total of 244 undergraduates from an introductory course in psychology at the University of Central Florida participated in this study. Of them, 14 participants had to be excluded after data screening (see Results section). The resulting population was made up of 138 women and 93 men. All participants were at least 18 years of age, and the median age for participants was 19 years. Although the overall stimuli were specifically geared toward a very particular aviation task which this population was likely not accustomed to, the goal was to study category learning rather than category representation. As such, category learning was best observed using populations previously unfamiliar to the categories, in a context free task, so that true learning could be observed (Ashby & Maddox, 2005).

### **Design**

The goal of this study was to bridge the gap between an applied perceptual training task and a model of category learning. To accomplish this, the study manipulated two variables in a perceptual training task directly: (a) complexity and (b) training method. This led to a 4 (simple stimulus, target stimulus, proximal stimulus, full environment stimulus) x 2 (single exposure vs. discrimination image display method) between-subjects design. A no-training control group was also included in the design to provide baseline comparisons on the transfer task.

## Experimental Task

In aviation, the approach and landing remains one of the most difficult phases of flight. Analysis of aviation accident and incident data reveal that controlled flight into terrain (CFIT) remains one of the most prevalent causes of fatality (Darby, 2006; Shappell & Wiegmann, 2003). Although there are numerous factors that can influence a pilot's decision process which can contribute to CFIT, the visual approach has consistently been identified as a specific maneuver that is difficult to train and can in some instances contribute to CFIT. The visual approach requires pilots to rely on their visual perception of the environment out-of-the-cockpit to keep safe separation from surrounding traffic and maintain a stable angle of approach (a.k.a., glideslope) to the landing surface. Perceptually, this requires a pilot to be able to recognize instances when they have deviated from recommended glideslope (i.e., am I too high or too low) or be aware of the perceptual differences they may encounter in non-standard approaches which operate outside of normal approach flight parameters. Unabated misperceptions in these instances can lead to unstable approaches which sometimes result in unsafe flying conditions, and in extremes can lead to CFIT. In terms of perceptual judgment, the visual approach offers a relevant task domain in which a replication of previous studies, which used simple stimuli to test the ATRIUM model, can be executed using a more applied task. The experimental task in this study was designed to closely resemble and in some respects mirror the design used in previous simple stimuli studies involving the ATRIUM model (Erickson & Kruschke, 1998). This was accomplished using a method in which there was a primary dimension in which a majority of items were categorized according to where it falls in relation to the rule boundary. A secondary

dimension was also presented in which, in select cases, the combination of primary and secondary dimension values generated an exception to the primary category rule.



Figure 2. Sample static visual approach image. Generated using Microsoft Flight Simulator: X (Microsoft 2006).

The task in this study required participants to make perceptual judgments of static images of visual approaches (Figure 2). Previous research using the visual approach task domain, have shown that non-pilot participants are capable of performing the task, also showing improvement on task performance (Curtis, Jentsch, & Maraj, 2009). The task goal was to accurately label displayed approaches into appropriate categories based on relevant cue dimensions in the environment. In order to accurately accomplish this, participants should have attended to two relevant cue dimensions. The primary dimension for the task was based on an individual's ability to judge differences in glideslopes. Glideslope was defined as the angle of descent resulting from the combination of altitude and ground distance from the landing surface (Figure 3). Generally in

aviation, a  $3^\circ$  glideslope is the recommended path of stable approach. Anything that falls above or below that path in normal circumstances could be labeled too high or too low. In its simplest conception, glideslope can be thought of as a distance judgment between two points in space (i.e., end of the runway and horizon). In normal flight operations, this dimension alone can be used to judge whether or not the approach is on a stable approach path. In non-standard approaches, the recommended glideslope may differ from  $3^\circ$ . In these cases a secondary source of information is necessary to accurately make perceptual judgments.

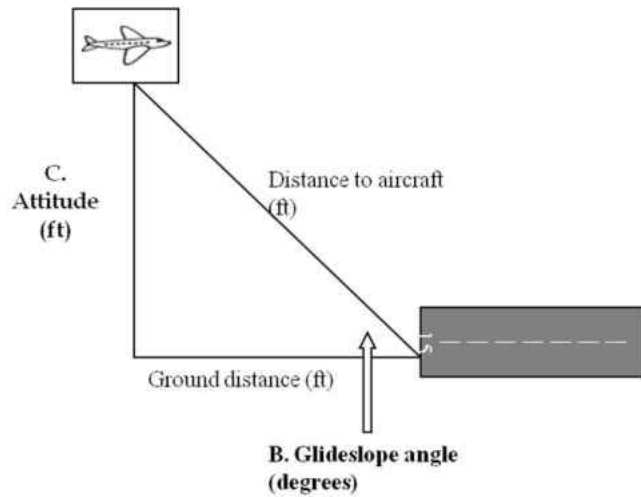


Figure 3. Geometric properties of glideslope.

The Precision Approach Path Indicator (PAPI) is a lighting system that provides a source of glideslope information to pilots. It is comprised of four lights which depending on the glideslope that an aircraft approaches will change combinations of red and white lights (Figure 4). In a true replication of an aviation task the PAPI provides a strong external cue to aid pilot

decision making in a visual approach. However, to maintain the primary/secondary dimension distinctions that more closely matches the previous ATRIUM category learning research, a slight modification of the true PAPI lighting configuration was used as the secondary dimension. Despite the modification described in more detail in the following section, the underlying rules that govern these two dimensions were preserved to maintain face validity of the aviation task.

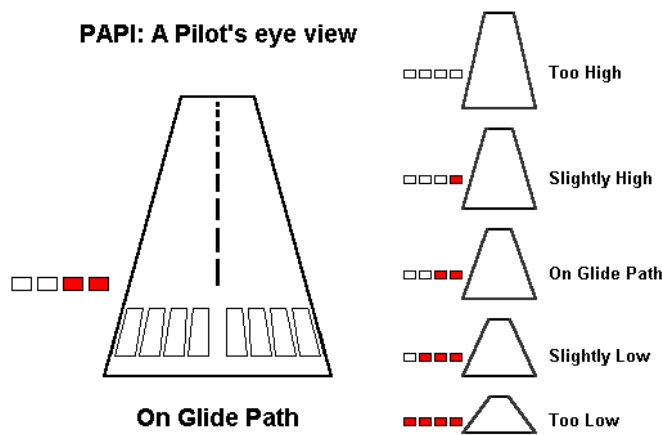


Figure 4. Precision Approach Path Indicator lighting configuration.

### Category Labels

Categories in which each visual approach image was labeled were based on the primary and secondary dimensions. In order to correctly label all images, participants had to attend to both. In order to preserve a context free task environment, category labels used in the task were nonsense words to reduce the potential for conceptual priming. The primary category distinction was made based on whether the visual approach image appeared above or below the primary dimension category boundary ( $3^\circ$  glideslope). Within these two categories there were five

possible glideslope values that differed in half-degree ( $0.5^\circ$ ) increments. All items that were above the primary dimension boundary were labeled *Spulch*. Items that are below the primary dimension were labeled *Trantac*.

The secondary, rule-exception, category labels were derived from a specific combination of primary and secondary values. As discussed above the secondary dimension was based on the PAPI decision aid common at many airports. To replicate the task from previous studies closely, yet maintain some level of face validity for the aviation task, the PAPI configurations were presented in a way that two of the four lights were blacked out and the remaining images were combinations of red and white lights commonly seen in an approach. The secondary dimension was comprised of 10 unique lighting combinations. In the training conditions, only two of these lighting conditions in combination with specific glideslope values produced the rule exception categories, non-standard steep or non-standard shallow approach. The non-standard steep item was labeled *Yarp* and the non-standard shallow approach item was labeled *Peltonic*.

#### Apparatus

##### *Administration*

This study was administered on 17" wide screen laptop computers. All study materials were presented using MediaLab and Direct RT research software.

*Stimuli*

All images were generated using images generated in Microsoft Flight Simulator X: Deluxe Edition (Microsoft, 2006), Team Performance Lab-Where Are You (TPL-WAY) software, and Adobe Photoshop. Microsoft flight simulator served as the platform from which the imagery was generated using still shots from visual approach conditions. In order to maintain image precision, the TPL-WAY software (Curtis, Schuster, Jentsch, Harper-Sciarini, and Swanson, 2008) was used to provide precise geo-global coordinates for aircraft positioning at real airport locations. In order to control airport and runway size in varying terrains in the complex stimuli condition, Adobe Photoshop was used to edit the still images created in Microsoft Flight Simulator. Additionally Adobe Photoshop was used to remove visual information from the images for the task isolation conditions.

Table 3. Matrix of category structure for relevant task dimensions in the perceptual training task. Rows represent stimulus values along the primary dimension, change in glideslope. Columns represent stimulus values along the secondary dimension, PAPI configuration. Training stimulus is designated by a cell containing either an H or an L depending on which side of the primary category boundary the stimulus falls. The two exception categories are labeled EH and EL.

<b>Glideslope Change</b>	5.5		H				H			
	5					H				
	4.5	H	EH				H		H	
	4				H					
	3.5		H				H			
	2.5				L				L	
	2					L				
	1.5	L		L				EL	L	
	1				L					
	.5			L				L		
	WBWB	BWWB	WWBB	BBWR	BWBR	WBRB	WRBB	BBRR	BRRB	BRBR
	<b>PAPI Configuration</b>									

## Training

Training was comprised of a selection of 20 images from the matrix of possible primary and secondary cue dimension combinations (Table 3). Of these 20 images, 18 could be categorized in one of the primary category labels (9 as “high approach” and 9 as “low approach”); the remaining two images were exceptions to the primary rule that each had a category label (non-standard high or non-standard low). These 20 images were presented in multiple training blocks. Within each training block, all of the rule images were presented once. The items labeled EH and EL in the stimulus matrix represented exceptions to the primary categorization rule. These were presented three times in each training block. To assess the predictive effectiveness of the ATRIUM model of category learning to an applied task environment and to investigate the effectiveness of different image exposure methods, there were two manipulations to the training.

### *Training Method*

The training method manipulation focused on the way in which the stimuli were displayed during training and the response task they were asked to perform. There were two conditions, single image exposure training and two image discrimination training. In addition to a difference in number of images exposed per training item, each condition asked participants to respond in different ways.



### *Single-Image Exposure*

In the single-image exposure condition, 24 training images were presented, one at a time, to the trainee within each training block. The order of presentation was randomized across all training blocks. The task in this condition was to categorize the displayed stimuli (spulch, trantac, yarp, peltonic) based on the critical cue dimensions that determined category membership.

### *Two-Image Discrimination*





In the two-image discrimination condition, 24 side-by-side image pairs were presented to the participant. The task was to determine if the two images presented were the same or different based on critical cues which determine category membership in the environment. For this study, the discrimination task centered on training individuals to gauge differences in glideslope and lighting variations in the displayed visual approach image that dictated category membership. To prevent ceiling and floor effects for the task, the selected stimulus discriminations were counterbalanced for within category and between category distances. Larger differences within categories and smaller distances between categories were presumed to be harder to correctly discriminate.

### *Image Complexity*

The second manipulation in this study was that of image complexity. To investigate the impact of irrelevant, but environmentally present, clutter on category learning, four training conditions were developed. Increase in task complexity across conditions was defined in terms of

the proximity and number of irrelevant stimulus manipulations to the task relevant stimuli (see Table 4). This provided a continuum from simple stimulus training where there were no irrelevant pieces of information displayed to full environment stimulus training in which the desired skill was embedded in a cluttered environment of varying irrelevant information that should have been suppressed to successfully complete the task. The following section provides a brief description of each.

Table 4. Complexity manipulation table. Illustrates the irrelevant stimulus dimensions for each level of complexity.

Complexity Condition	Simple Stimulus	Target Stimulus	Proximal Stimulus	Full Environment Stimulus
Image sample				
Irrelevant cue	<ul style="list-style-type: none"> <li>- N/A</li> </ul>	<p>Runway Size</p> <ul style="list-style-type: none"> <li>- Long/Narrow Runway</li> <li>- Short/Wide Runway</li> </ul>	<p>Runway size</p> <ul style="list-style-type: none"> <li>- Long/Narrow Runway</li> <li>- Short/Wide Runway</li> </ul> <p>Airport size (# of runways)</p> <ul style="list-style-type: none"> <li>- One runway airport</li> <li>- Multiple Runway Airport</li> </ul>	<p>Runway size</p> <ul style="list-style-type: none"> <li>- Long/Narrow Runway</li> <li>- Short/Wide Runway</li> </ul> <p>Airport size (# of runways)</p> <ul style="list-style-type: none"> <li>- One runway airport</li> <li>- Multiple Runway Airport</li> </ul> <p>Terrain density</p> <ul style="list-style-type: none"> <li>- Desert terrain</li> <li>- City Terrain</li> </ul>

### *Simple Stimulus Training*

The simple stimulus condition was intended to isolate the target skill from the environment. This provided practice without distraction from other stimuli in the environment. For this task, a simple stimulus consisted of an image of a vertical line which represented a very simple depiction of a runway, a series of four dots to the left of the vertical line, and a horizontal line that spanned the length of the image that represented the horizon line. No additional perceptual information was provided in this condition. By isolating the critical dimensions from a complex perceptual environment this condition represented a pure task isolation condition.

### *Target Stimulus*

The target stimulus condition introduced an irrelevant dimension to the categorization task: runway size. While the dot configuration and horizon variables remained the same, the line representing the runway was replaced with an image of a runway which varied in size. This runway size manipulation was intended to replicate a commonly documented distraction pilots encounter when estimating glideslope in a visual approach known as form ratio. Form ratio was defined as the ratio of the length and width of a runway. In some cases where runways are at extreme dimensions (e.g. short/wide or long/narrow) pilots have a tendency to over or under estimate their position in space. The irrelevant dimension in this case was a direct manipulation of the appearance of one of the target dimensions in the categorization task.

### *Proximal Stimulus*

The proximal stimulus condition served to broaden the scope of visual clutter to features in the environment that were in close proximity to the target cues in the environment. In addition to the irrelevant manipulation of runway size, this condition presented additional irrelevant information of the airport surrounding the target runway. This manipulation was intended to replicate conditions where the size of the airport may impact the ability of a pilot to locate the appropriate landing surface. Small airports consisting of only one runway (the target runway) may have yielded far different results than a large airport with many runways running either in parallel or crossing each other.

### *Full Environment Stimulus*

Building off of the proximal stimulus condition, the full environment stimulus condition included all irrelevant features previously discussed (i.e., runway size, airport size) in addition to broadening the terrain information to the entire image. This condition served to replicate the phenomena where the density of the surrounding environment terrain can influence how one perceives their position in space. The full environment stimulus condition represented the opposite end of the spectrum from the task isolation condition. Whereas the simple stimulus training condition represented a full isolation of the target skill (visual glideslope estimation) the full environment stimulus condition represented the embedding of the target skill into an environment cluttered with visual information that should have been suppressed to effectively make perceptual judgments.

### *Control Group*

In order to provide a baseline for performance on the transfer task, a no training control conditions was also included. Participants in this condition received only pre-training which was necessary to provide relevant information on how to perform the transfer task. Participants in this condition were given the transfer tasks immediately following pre-training.

### *Summary of Training*

The training task was designed to address the impact of training type and image complexity on response patterns and overall performance on a transfer categorization task. Given that there were a varied number of factors that impacted image content for individual training items, an effort was made to counterbalance each item variable. Participants were exposed equally to each variable instance in their respective training condition. Image variables like runway size, airport size, and terrain were all counterbalanced so that each occurred in as close to an equal number of condition appropriate items as possible across all training blocks. Due to anticipated response patterns of interest, a consistent but unequal number of items were selected on the primary (glideslope) and secondary (PAPI lighting) conditions (previously demonstrated in Table 3) for training. This intentional imbalance was intended to focus on the variable values that were anticipated to produce the most relevant response variance. Image sets were balanced across training blocks so that all participants in each training condition were exposed to the same images.

## Measures

### *Demographics*

The demographic questionnaire contained standard population demographics (e.g., gender, age, etc...), aviation experience, and video game experience questions. See Appendix A for the full questionnaire form.

### *Spatial Orientation*

The Guilford-Zimmerman spatial orientation (Guilford & Zimmerman, 1948) measure was used for this study. This specific measure was chosen due to previous research that showed that this measure was predictive of performance on similar aviation training tasks (Curtis, Jentsch, & Maraj, 2009). See Appendix B for the measure. The spatial orientation measure was intended as a covariate in analysis.

### *Skill Transfer*

To test the transfer impact of the training task, two skill transfer tasks were used. Both were 100-item single image categorization tests. Participants viewed a display in which a single image was shown and asked to categorize whether the image represented a too high, too low, non-standard steep or non-standard shallow approach based on the nonsense word labels associated with them. The first of the two transfer tasks was a simple stimulus transfer task. The images in this task were comprised of the entire pool of simple stimuli generated for the stimuli matrix listed in the design section of this paper. The second transfer was a suppression transfer task in which the images were comprised of items from the pool of possible full environment suppression images. The addition of irrelevant cue manipulations inflated the number of possible

variable combinations to 800. To maintain practicality of test time, using the same technique as described in the training section, a counterbalanced selection of 100 images from the full environment variable image pool were selected for use in the transfer test.

### Procedure

Upon arrival, each participant was randomly assigned to one of two training method conditions and one of four complexity conditions or to a no training control group. Participants were seated in front of a laptop computer where they were asked to read and verbally agree for informed consent. After that, participants were asked to answer a series of demographic questions and completed a test of spatial ability. Prior to beginning the training participants received a brief description of the categorization task they were asked to perform and given a short test that indicated comprehension of the category labels. Participants then began the training. They received 16 blocks of a perceptual training task that matched the training method and complexity condition they were assigned. Each perceptual image comparison was presented for a maximum of 10 seconds, and knowledge of results feedback was provided following each response. In the training, participants were asked to categorize the pair of images by pressing designated keys on the keyboard. Failure to respond within 10 seconds resulted in an incorrect response. Participants were given a short break half way through the training blocks. At the end of the discrimination training, a skill transfer post-test consisting of 200 images was administered. At the conclusion of the post-test, participants were debriefed and dismissed.



## CHAPTER 4: RESULTS

For this study, analyses consisted of a series of mixed factorial ANOVAs based on hypothesized relationships for variables relevant to performance on the simple and complex transfer task. Descriptive statistics are presented first, and what follows is a more detailed description of the analyses for each hypothesis.

### Descriptive Data

Prior to analysis, the data were screened. Participants with an average combined mean performance score of 25% or less for the final three training blocks (i.e., Block 14, 15 and 16) were excluded from analysis. Only eight participants had to be excluded using this criterion, and, in each case, patterns of responses clearly indicated that these were participants who had stopped trying and instead were merely “clicking through” each stimulus. In addition, any participants who did not receive one of the performance measures, or had computer malfunctions during their training session were also excluded from analysis. Together, application of the two criteria resulted in the exclusion of a total of 13 participants from the final analysis. Among the remaining participants, the number of participants per condition was nearly evenly distributed (see Table 5).

Within the population, only 14 participants reported having any flight experience, due to the low number of flight hours reported by these participants, no further distinction was made between experienced and inexperienced participants in this data set. Mean, standard deviation and inter-correlations between important variables are outlined in Table 6.

Table 5. Population frequency for participant conditions

Condition	Simple	Runway	Airport	Full	Training Type N
<b>Exposure</b>	25	26	25	25	101
<b>Discrimination</b>	27	24	25	25	101
<b>Complexity Condition N</b>	52	50	50	50	202
<b>Control Group</b>					29
				<b>Overall N</b>	231

Table 6. Means, standard deviations and inter-correlations of primary study variables.

	Exposure Means	Discrimination Means	Control Means	IC	B1-O	B1-E	B1-OG	B16-O	B16-E	B16-OG	ST	SE	PCES	SCES	FT	FE	PCEF	SCEF	GZ-Spat. Ori.
Image Complex.				-															
B1-Overall	.388 (.135)	.587 (.123)	--	-.006	-														
B1-Exceptions	.394 (.249)	.636 (.209)	--	-.173*	.729**	-													
B1-Overgen	.401 (.236)	.441 (.298)	--	.239**	-.098	-.547**	-												
B16-Overall	.602 (.189)	.644 (.131)	--	-.158*	.256**	.323**	-.049	-											
B16-Exceptions	.417 (.319)	.663 (.209)	--	-.186*	.389**	.421**	-.095	.607**	-										
B16-Overgen.	.477 (.297)	.453 (.339)	--	.191**	-.088	-.177*	.151*	-.299**	-.734**	-									
Simple Transfer	.528 (.247)	.271 (.151)	.249 (.121)	-.048	-.305*	-.122	-.032	.338**	-.079	-.002	-								
Simple Except.	31.58 (17.80)	51.039 (8.867)	51.58 (7.27)	-.045	.349**	.163*	.037	-.356**	.150*	-.060	-.786**	-							
Prim.Exc Simp	.216 (.191)	.292 (.199)	.322 (.177)	-.066	.166*	.061	.030	-.145*	.065	-.037	-.058	.470**	-						
Sec.Exc. Simp.	.250 (.200)	.285 (.183)	.278 (.198)	.001	.160*	.123	.016	.025	.251**	-.178*	-.263**	.245**	-.129*	-					

Full Transfer	.605 (.184)	.345 (.126)	.379 (.150)	.120	-.249*	-.194**	.020	.316**	-.132	.059	.597**	-.673**	-.220**	-.111	-				
Full Except.	29.94 (16.72)	52.347 (7.669)	49.24 (11.52)	-.127	.402**	.267**	-.027	-.200**	.227**	-.047	-.559**	.732**	.370**	.105	-.828**	-			
Prim. Exc. Full	.235 (.192)	.402 (.227)	.393 (.235)	-.023	.280**	.139*	.098	-.019	.207**	-.072	-.244**	.433**	.367**	.000	-.325**	.650**	-		
Sec. Exc Full	.215 (.152)	.266 (.159)	.291 (.166)	-.199*	.151*	.177**	-.099	.159*	.308**	-.170*	-.134*	.169*	-.057	.509**	-.283**	.251**	-.063	-	
Guilford-Zim. Spat. Ori	19.17 (8.91)	21.21 (10.17)	--	.184**	.189**	.261**	-.065	.135	.235**	-.177*	.051	.017	-.059	.144*	.039	-.002	.139*	.018	-
Mean (SD)				22.17 (1.34)	.49 (.16)	.51 (.26)	.42 (.27)	.62 (.16)	.54 (.30)	.46 (.32)	.38 (.23)	42.60 (16.53)	.26 (.20)	.27 (.19)	.46 (.20)	42.16 (16.75)	.33 (.23)	.25 (.16)	18.71 (10.63)

\*p < .05, \*\*p < .01

Note: Image Complexity is the between subjects manipulation of training image complexity. B1-Overall and B16- Overall represent the percentage of correct overall responses within a specific training block (i.e. B1= training block 1, B16= training block 16). B1-Exception and B16-Exception represent the percentage exception items that are correctly identified within a specific training block. B1-Overgen and B16-Overgen represent the percentage of exception items that are incorrectly labeled according to the corresponding rule. Simple Transfer and Full Transfer are the percentage of correct identified items for the respective transfer test. Simple Except and Full Except are the total number of exceptions responses for the respective transfer test. Prim Exec Simp and Prim Exec Full are the percentage of transfer task items that have the same primary task cue that are labeled as exceptions. Sec. Exc. Simp and Sec. Exc. Full are the percentage of transfer task items that have the same secondary task cue that are labeled as exceptions. Guilford-Zim. Spat. Orientation is the number of correct items identified on the Guilford-Zimmerman Spatial Orientation measure. Means and standard deviation are provided for the training type condition.

### Hypothesis 1a-d: Effects of Stimulus Complexity on Model Fit

Hypothesis 1 proposed that as stimulus complexity increased, response patterns that predict fit in the ATRIUM model would hold. To compare model fit across conditions, four variables associated with response patterns that would predict model fit were tested: overgeneralization, exception response in transfer, correct exceptions in training, and transfer cue response pattern. Each sub-hypothesis represents analysis of these variables individually. Since previous tests of the ATRIUM model used exposure training, exposure and discrimination training conditions were analyzed separately. This allowed tests of (a) whether the model held in a more complex replication and extension of the ATRIUM study to naturalistic stimuli, and (b) whether it would transfer to a training method previously untested with this model, under the same conditions. Since the discrimination training group performed equally poorly on the transfer tasks as a no training control group, caution was exercised in drawing conclusions in respect to training effectiveness response patterns that emerged from any measure using the transfer task in Hypothesis 1c and 1d.

*Hypothesis 1a: As training stimulus complexity increases, participant reliance on rule-based response strategies will increase such that overgeneralization will be evident in training for more trials than in simple stimulus conditions.*

*Exposure Condition.* Hypothesis 1a proposed to find a relationship between percentages of exception items in each training block that were overgeneralized. A mixed factorial ANOVA comparing overgeneralization scores across four complexity conditions and 16 training blocks

was conducted. A significant within-subjects effect was present for overgeneralization  $F(15, 1455) = 3.021, p < .001, \text{Partial } \eta^2 = .263$ , and a significant between-subjects effect for complexity  $F(3, 97) = 2.448, p = .068, \text{Partial } \eta^2 = .070$ , were observed. There was no significant interaction observed for overgeneralization and complexity on the analysis. Pairwise comparisons were conducted on Training Block 1 to test for initial differences between conditions. Pairwise comparisons were also conducted on Training Blocks 14, 15, and 16 to see if group differences emerged in the final training blocks. This rationale for comparison holds for all remaining hypotheses throughout the results section (i.e., Hypothesis 1c) which involve within subjects comparisons across training blocks. For Training Block 1, the percentage of overgeneralized exceptions in the simple image condition ( $M = .320, SD = .198$ ) was significantly less than the proximal image condition ( $M = .447, SD = .203$ ), and the full image conditions ( $M = .447, SD = .258$ ).

With respect to the final blocks of the training session, there were no significant differences for Training Block 14. In Block 15, the simple image condition ( $M = .340, SD = .345$ ) had significantly less overgeneralized responses than the target image condition ( $M = .577, SD = .321$ ), the proximal image condition ( $M = .513, SD = .220$ ), and the full environment image condition ( $M = .527, SD = .271$ ). Similarly, in Block 16, percentage of overgeneralization in the simple image condition ( $M = .327, SD = .328$ ) was significantly less than in the target image condition ( $M = .564, SD = .306$ ), in the proximal image condition ( $M = .507, SD = .212$ ), and in the full environment image condition ( $M = .507, SD = .291$ ). The differences between groups in

Block 1 suggest that the simple presence of distracting visual information may be enough to affect overgeneralization behavior. Similar patterns in the final training blocks, in spite of lack of significant differences between the three conditions that contained distracting information, largely support Hypothesis 1a. The pattern of responses across training Blocks 1, 14, 15, and 16 (Figure 5), illustrates a trend of increasing overgeneralization in the three conditions which included distracting additional information.

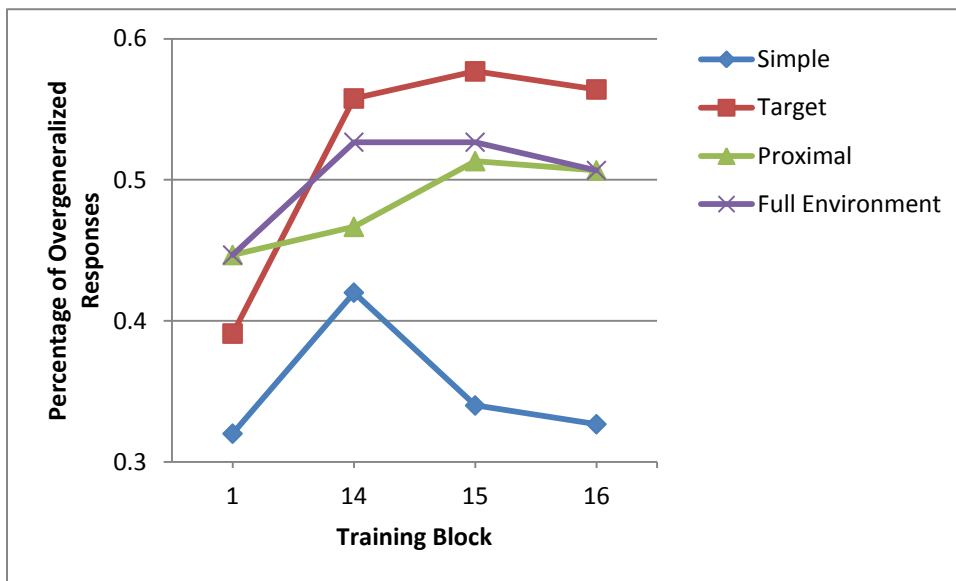


Figure 5. Percentage of overgeneralized responses for training Block 1, 14,15,and 16 by training image complexity condition for the exposure training condition.

*Discrimination Condition.* Analysis of overgeneralization for the discrimination training condition produced a significant main effect for training image complexity  $F(3, 97) = 5.704, p < .01, \text{Partial } \eta^2 = .150$ . Pairwise comparisons revealed that participants that were in the simple

image condition ( $M = .377, SD = .182$ ) had a significantly lower percentage of overgeneralized responses than the target ( $M = .486, SD = .177$ ), proximal ( $M = .561, SD = .176$ ) and full environment ( $M = .538, SD = .167$ ) conditions. Although there were no significant within-subjects effects for overgeneralization across training blocks, the between subject effect was similar to the between subjects effect in the exposure condition, suggesting that the pattern of overgeneralization across complexity groups is consistent between exposure and discrimination training.

*Hypothesis 1b: As training stimulus complexity increases, participant reliance on rule-based response strategies will increase such that a lower proportion of exception training stimuli will be labeled correctly.*

*Exposure Condition.* Another mixed factorial ANOVA was conducted to address Hypothesis 1b. The comparison involved within subjects variable, correct exceptions in training, and between subjects variable image complexity. For the exposure training condition, there were significant main effects for correct exceptions in training  $F(15, 1545) = 1.782, p < .05, Partial Eta^2 = .017$ , and image complexity  $F(1,103) = 2.329, p = .079, Partial Eta^2 = .064$ . Additionally there was a significant interaction between correct exceptions and image complexity  $F(45, 1545) = 1.878, p < .001, Partial Eta^2 = .052$ . Pairwise comparisons were conducted on training blocks 1, 14, 15, and 16. For Training Block 1, there were no significant differences between conditions. Analysis of Block 14 revealed that the simple image condition ( $M = .527, SD = .369$ ) had significantly higher percentage of correct exceptions than the proximal image condition ( $M =$

.373,  $SD = .260$ ), and the full environment image condition ( $M = .280$ ,  $SD = .219$ ). Analysis of Block 15 revealed that the simple image condition ( $M = .607$ ,  $SD = .378$ ) had significantly higher percentage of correct exceptions responses than the target image condition ( $M = .339$ ,  $SD = .348$ ), the proximal image condition ( $M = .360$ ,  $SD = .224$ ) and the full environment image condition ( $M = .360$ ,  $SD = .283$ ). Similarly, in Block 16, the simple image condition ( $M = .613$ ,  $SD = .359$ ) showed significantly higher performance than the target image condition ( $M = .333$ ,  $SD = .319$ ), proximal image condition ( $M = .367$ ,  $SD = .215$ ), and the full environment image condition ( $M = .360$ ,  $SD = .295$ ). Figure 6 further illustrates the pattern for training Block 1, 14, 15, and 16. Although there were no differences between the target, proximal, and full environment conditions, the significant difference between each of these conditions and the simple conditions largely supported Hypothesis 1b. Furthermore, the pattern of responses for the exposure condition served as a logical contrast to the results of Hypothesis 1a. It makes sense that as the tendency to overgeneralize decreased that the number of correctly labeled exceptions increased.



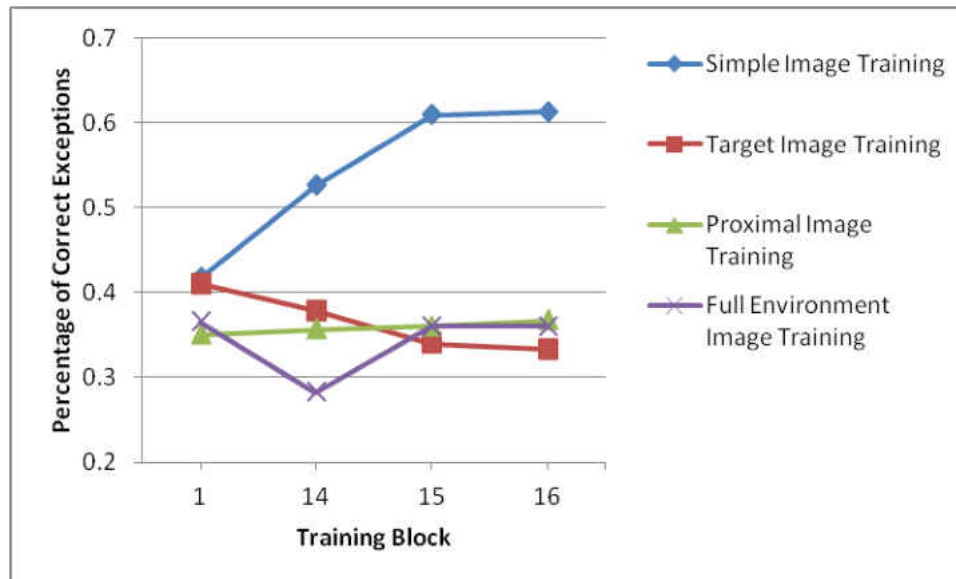


Figure 6. Percentage of correct exception responses across training block by image complexity training condition in exposure training condition.

*Discrimination Condition.* Analysis of the discrimination training group produced a significant between subjects main effect for image complexity  $F(3, 97) = 3.805, p < .05, Partial \eta^2 = .105$ . One tailed simple effects analysis showed that the simple image condition ( $M = .703, SD = .130$ ) was significantly higher than the target image condition ( $M = .642, SD = .642$ ), the proximal image condition ( $M = .602, SD = .098$ ), and the full environment image condition ( $M = .627, SD = .117$ ) across training conditions. Since there was no main effect for training block or no significant interaction between training block and image complexity, no further comparisons were conducted. Similar to Hypothesis 1a, there was a significant effect for complexity however there was no significant change within groups across the training block. The pattern of these sub-hypotheses supports the model predictions based on increased complexity, but did not support any improvements associated with the discrimination training.

*Hypothesis 1c: As training stimulus complexity increases, participant reliance on rule-based response strategies will increase such that less overall exception responses will be given in the transfer task.*

*Exposure Condition.* A second variable predictive of model fit, total exception responses in transfer, were analyzed in Hypothesis 1c across four image complexity conditions using a mixed factorial ANOVA. For the exposure training condition, there was no significant effect for either main effects (total number of exception responses or image complexity), however, there was a significant interaction for total exception responses and image complexity  $F(3,97) = 8.373$ ,  $p < .001$ ,  $Partial\ Eta^2 = .206$ . Planned comparisons show that there are significantly less exceptions in the simple image complexity condition in the simple transfer ( $M = 24.3$ ,  $SD = 19.91$ ) than in the full environment transfer condition ( $M = 32.8$ ,  $SD = 19.69$ ). In addition, for the proximal image complexity condition there were significantly more exception responses in the simple transfer ( $M = 39.16$ ,  $SD = 14.82$ ) than in the full environment transfer ( $M = 32.96$ ,  $SD = 11.58$ ). A similar pattern was also evident in the full environment complexity condition with the simple transfer ( $M = 34.96$ ,  $SD = 16.518$ ) and full transfer condition ( $M = 26.96$ ,  $SD = 18.160$ ). Although this does not fully follow the hypothesized relationship, the results suggested that the more familiar transfer stimuli result in less exception responses. Figure 7 further illustrates the nature of these relationships.

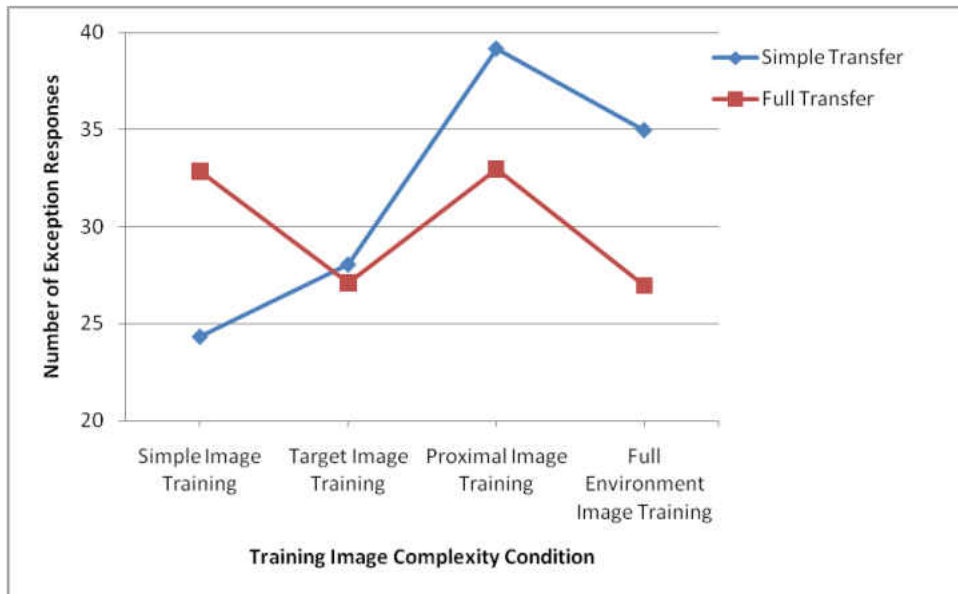


Figure 7. Total number of exception responses by transfer type across training complexity condition for exposure training group.

*Discrimination Training.* For the discrimination training condition, there were no significant main effects or interactions. This is not surprising considering the low overall transfer task performance scores for the discrimination training condition discussed in hypothesis 2b.

*Hypothesis 1d:* As complexity increases greater reliance on rule-based processes will impact dimensional attention in the transfer task such that a higher proportion of exception responses will occur on the primary (rule) dimension of exceptions than on the secondary (exception) dimension.

*Exposure Condition.* Analysis of Hypothesis 1d was conducted using a mixed factorial design involving three factors. The between subjects factor was complexity and there were two

within subjects factors, exception response pattern and transfer task complexity. In the exposure training condition, there were no significant main effects, but there were three interaction effects: First, there was a two-way interaction for transfer task and training image complexity  $F(3, 104) = 2.754, p < .05, \text{Partial } \eta^2 = .074$  (Figure 8). Second, there was also a two way interaction for transfer task and exception response pattern  $F(1,104) = 3.928, p = .05, \text{Partial } \eta^2 = .036$  (Figure 9). Planned comparison showed a significantly lower percentage of primary exception responses in the simple transfer task ( $M = .262, SD = .196$ ) than the full environment transfer task ( $M = .328, SD = .228$ ). There were also significantly more primary exception responses ( $M = .328, SD = .228$ ) than secondary exception responses ( $M = .247, SD = .159$ ) in the full environment transfer task. These effects supported the assertion in Hypothesis 1d that increased image complexity would result in a higher proportion of primary cue exception responses.

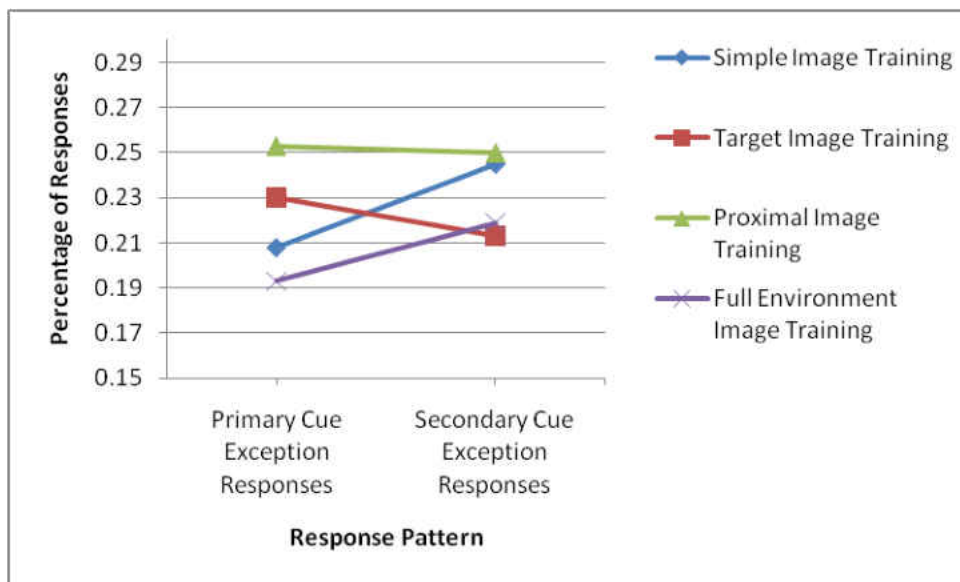


Figure 8. Illustration of interaction between exception response pattern and training image complexity.

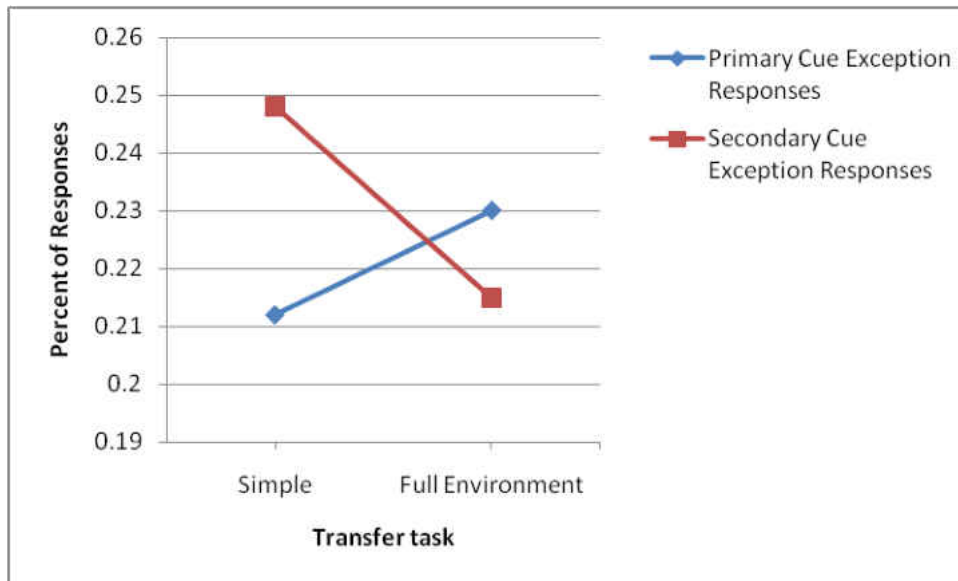


Figure 9. Illustration of interaction between transfer task and exception response pattern.

Third, and finally, there was a three-way interaction between transfer task, exception response pattern, and training image complexity  $F(3,104) = 3.366, p < .05, Partial Eta^2 = .088$ . Pairwise comparisons on the three-way interaction showed significant differences between secondary cue exception responses in the simple and full transfer conditions for the simple image complexity condition ( $M_s = .193$  vs.  $M_f = .282$ ), the proximal condition ( $M_s = .302$  vs.  $M_f = .204$ ), and the full environment condition ( $M_s = .276$  vs.  $M_f = .178$ ). This provided indication that responses based on the secondary cue increased in the transfer task which was dissimilar to the training stimuli.

*Discrimination Condition.* For the discrimination training condition, there were within-subjects main effects for transfer task  $F(1, 97) = 10.149, p < .01, Partial Eta^2 = .095$ , exception

response pattern  $F(1, 97) = 9.568, p < .01, \text{Partial } \eta^2 = .090$ , and a significant interaction between transfer task and exception response pattern  $F(1, 97) = 13.186, p < .01, \text{Partial } \eta^2 = .120$ . There were no significant between-subjects effects for complexity. Pairwise comparisons showed that participants provided significantly more primary cue exception responses ( $M = .403, SD = .227$ ) than secondary cue exception responses ( $M = .266, SD = .159$ ) in the full environment transfer. There were no significant differences in the simple transfer task. Due to low average means in Hypothesis 2b for transfer task performance, no further interpretation of these results for the discrimination task condition were provided.

#### Hypothesis 2a-b: Effect of Training Method on Category Learning

Hypothesis 2 proposed to address the relationship training method has with performance of a perceptual training task. Two indicators of training effectiveness, time to proficiency and overall transfer accuracy were hypothesized to provide an accurate picture of this. Analyses of these were represented in the two sub-hypotheses associated with this research question.

*Hypothesis 2a: By increasing the number of images per exposure, the discrimination training will result in reduced time to reach training proficiency.*

For Hypothesis 2a, a two factor mixed factorial ANOVA was used. The within subjects factor was percent correct per training block, and the between subjects factor was training type. There was a significant main effects for both training performance  $F(15, 3000) = 28.216, p < .001, \text{Partial } \eta^2 = .132$ , and training type  $F(1,200) = 67.593, p < .001, \text{Partial } \eta^2 = .258$ .

Planned comparisons indicated that the exposure condition ( $M = .510$ ,  $SD = .121$ ) had a significantly lower overall performance mean in training than the discrimination training ( $M = .628$ ,  $SD = .079$ ). In addition, there was a significant interaction for training performance by training type  $F(15, 3000) = 12.817$ ,  $p < .001$ ,  $Partial \eta^2 = .064$ . Although the discrimination training condition produced better scores there was only small performance improvement across training blocks. The pattern of results showed a more pronounced learning curve in the exposure condition. Although this seemed to support the hypothesis that the discrimination would reach a level of proficiency sooner, these results considered in combination with the outcomes on Hypothesis 1a and 1b seem to suggest that there was minimal learning in the discrimination, lending support to the notion that the two task training conditions were not similar enough for comparison. Although the exposure group never reached a learning plateau the performance increase across training blocks (Figure 10) indicated that participants learned as they progressed through training.

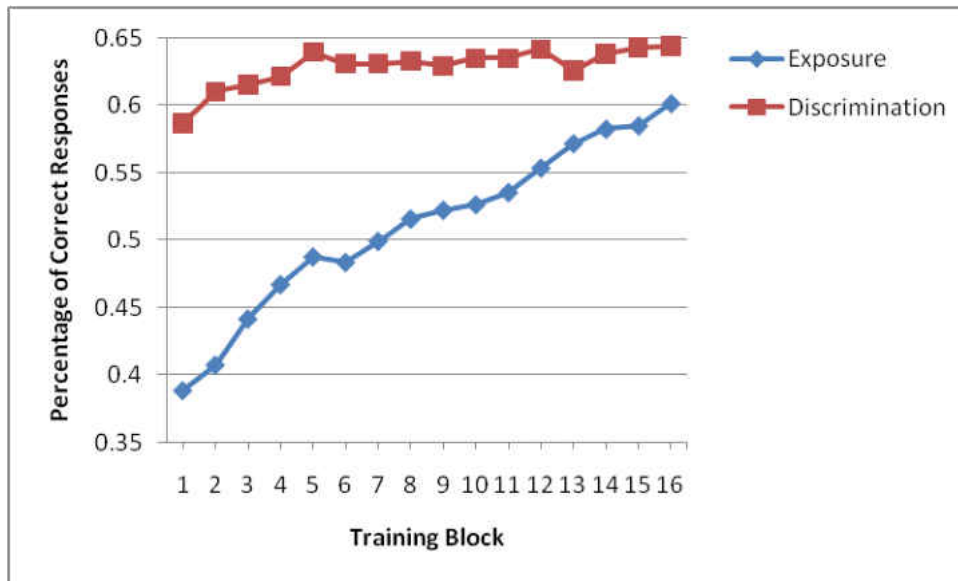


Figure 10. Overall performance across training between training conditions.

*Hypothesis 2b: Two image discrimination training elicits explicit practice with both exemplar and rule-based categorization strategies such that overall accuracy will be better on the transfer task than in the single exposure training condition.*

To analyze Hypothesis 2b, a mixed factorial ANOVA was used. Training type was the between-subjects variable, and transfer performance was the within subjects variable. There was a main effect for transfer task  $F(1, 228) = 37.457, p < .001, \text{Partial } \eta^2 = .141$ . Planned comparisons indicated better performance on the full environment transfer ( $M = .464, SD = .201$ ) than in the simple transfer task ( $M = .383, SD = .235$ ) across all training type conditions. More informative, there was a main effect for training type  $F(2, 228) = 86.808, p < .001, \text{Partial } \eta^2 = .432$ . Planned comparisons indicated significantly better performance in the exposure training



condition ( $M = .567, SD = .184$ ) than in the discrimination training condition ( $M = .308, SD = .114$ ) and a control group who received no training ( $M = .314, SD = .093$ ). There was no difference between the discrimination and control groups. There was no interaction for transfer by training type. Figure 11 illustrates the relationship of these variables.

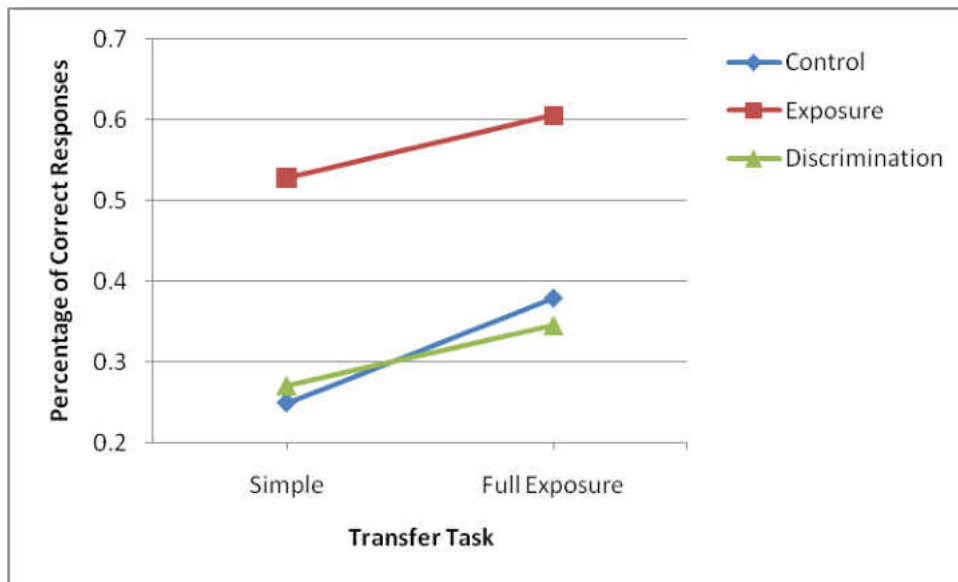


Figure 11. Transfer task performance between training conditions.

### Hypothesis 3a-b: Moderating Effect of Training Method on Training Stimulus Complexity

The final hypothesis proposed that training type would have a moderating effect on stimulus complexity in training and transfer. The combination of low transfer performance outcomes discussed in Hypothesis 2b and limited evidence for performance improvement in training outlined in Hypothesis 1a, 1b, and 2a suggested, that there were fundamental differences between the discrimination training and exposure training that made meaningful comparison and

interpretation of the two very difficult, if not impossible. The hypotheses that predicted improved performance in discrimination training over exposure training did not hold and, as a result of the condition performing no better than the control group, made further comparisons between the two training conditions predictable and ultimately of limited additional information. In spite of the lack of compelling comparison between training groups, it was determined that analysis of performance of the exposure training group in transfer revealed would be informative. Further description of this analysis is outlined below.

*Exposure condition performance in transfer.* A mixed factorial ANOVA was used to compare the between subjects variable training image complexity and the within subjects variable transfer task complexity. There was a significant main effect for transfer task complexity  $F(1, 97) = 17.968, p < .01, \text{Partial } \eta^2 = .156$ . Pairwise comparisons indicated significantly better overall transfer performance in the simple ( $M = .637, SD = .206$ ) and target ( $M = .618, SD = .184$ ) image complexity conditions than in the more complex proximal ( $M = .508, SD = .159$ ) and full environment ( $M = .502, SD = .151$ ) image complexity conditions. There was also significant main effect for training image complexity  $F(3, 97) = 4.116, p < .01, \text{Partial } \eta^2 = .113$ . Pairwise comparisons indicated better overall performance in the full transfer condition ( $M = .605, SD = .184$ ) than the simple transfer condition ( $M = .528, SD = .247$ ). Most interestingly, there was a significant interaction for training image complexity and transfer task complexity  $F(3, 97) = 20.208, p < .01, \text{Partial } \eta^2 = .385$  (Figure 12). The pattern shows that while transfer performance did not appear to be impacted by training image complexity in the

full environment transfer condition, there was a decrease in performance as complexity increased in the simple transfer task.

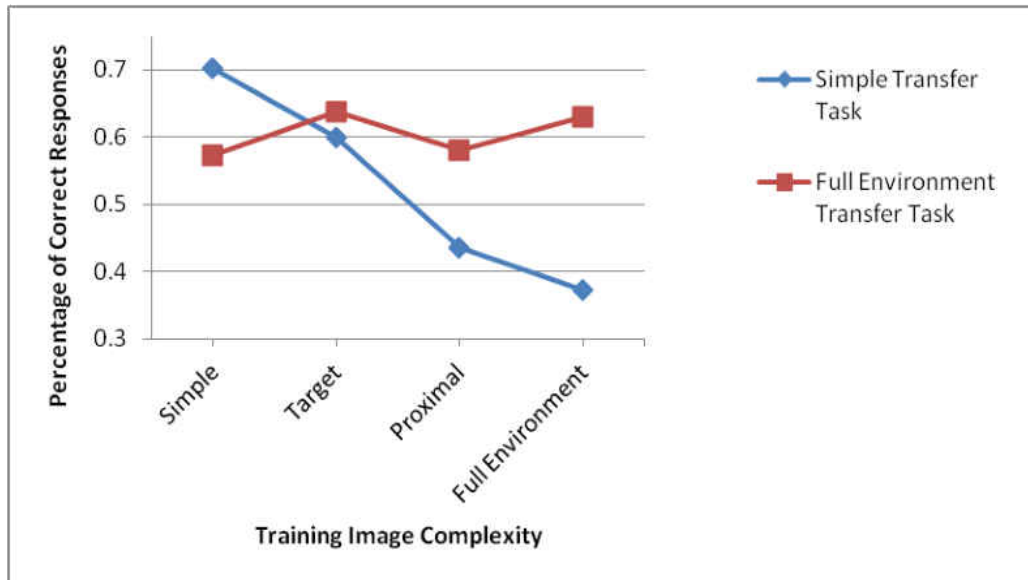


Figure 12. Percentage of correct responses in transfer tasks across training image complexity condition.

## CHAPTER 5: DISCUSSION

To reiterate what was previously stated, the goal of this research was twofold: 1) to validate a model of category learning using a real world task application and 2) to test the training effectiveness of a method of perceptual training. These goals were tested using a perceptual judgment task, specifically designed for a specific aviation context. The aviation task was adapted for a predominately novice population, in order to better test the category learning theory without preexisting biases. As will be discussed in this chapter, the study served as an important stepping stone for addressing a number of theoretical and applied questions related to perceptual judgments and training techniques. Ultimately, the findings from this study provided evidence for the extension of a model of category learning into more applied task domains, the importance of context when training with an unfamiliar perceptual task, and the benefit of task isolation in perceptual training.

Before engaging in specific discussion on the outcome of each hypothesis, an important outcome of the analysis should be noted. While participants in the discrimination training manipulation performed better than participants in the exposure training across all 16 training blocks, they failed to perform better than participants in a no training control group in the transfer tasks. The unexpected outcome reflected a consistent pattern of results, which did not support any hypotheses predicting overall performance in the discrimination condition. As a result, there are limitations to any discussion regarding the second goal of the research. As such, the remainder of the discussion will focus predominantly on the exposure training group. In spite

of the unexpected outcome, the failure of the discrimination training to transfer had important implications for the distinction between pure perceptual and category learning which will be discussed in this chapter.

The chapter is organized as follows. First is a summary of the results by hypothesis. A discussion on the theoretical implications of the research is next, which is followed by a discussion on the practical implications of the research. Study limitations and future research questions are addressed, followed by a closing with concluding remarks about the research.

## Hypothesis Discussion

### *Hypothesis 1*

The focus of Hypothesis 1 was to investigate how well response predictions of the ATRIUM model would hold up if the complexity of the trained image increased. To answer this question there were three criteria which would dictate whether or not the model extends. They were: 1) the simple image condition responses would follow predicted patterns of the original ATRIUM task, 2) adding complexity in training would result in more rule based responses, and 3) increasing complexity would magnify the rule based response effect.

The model was originally tested using exposure training with simple stimuli. The first objective of the study was to confirm that similar response patterns would be observed using an analogous perceptual task in a more applied task environment. The exposure training condition receiving the simple complexity images in training followed a pattern of response similar to previous ATRIUM studies (Erickson & Kruschke, 2001; Erickson & Kruschke, 1998). Through

the first 8 training sessions, there was a gradual increase in overgeneralization response and a relatively unchanged percentage of correct exceptions. In the final 8 training blocks, overgeneralization decreased from 50% to 32% of responses, and correct exception responses increased to 60% correct by the final training block. The diverging patterns of overgeneralization and correct exceptions followed predictions of the model (See Appendix D). As a result, the simple image condition was determined an acceptable proxy for the previous box-and-line ATRIUM task.

After establishing the simple image exposure condition closely followed model predictions, the next step was to compare results patterns across complexity conditions. The primary objective was to address training and transfer performance differences between using a simple image with only relevant visual information, and an image which incorporates all of the complexity of a real world task. Parsing the comparison into just these two conditions would have provided useful comparisons, but when considering model fit, would not have provided information relating to the additive effect that multiple visual distracters in combination may have. To more effectively pinpoint the effect increasing complexity as a function of distracting visual cues has on performance, two intermediary complexity conditions were also used in training. By providing a qualitatively conceptualized set of image complexity training conditions in a stepping function from simple to full complexity, it was easier to identify at what point, and to what degree, distracting information results in performance decrements.

The overall hypothesis was that there would be observable differences between all conditions as complexity increased. In the final training blocks, where it is logical to expect that differences between the groups would be most pronounced, there were no differences between any of the increased complexity conditions (i.e., target, proximal, full environment images) for overgeneralization or correct exception responses. In the final training conditions, overgeneralization was 16% lower in the simple complexity condition while also correctly identifying 17% more exception items than the three complexity conditions that had distracting information. The mere addition of one distracting visual cue in the training images was enough to disrupt category learning patterns as predicted by the ATRIUM model, and in relation to concepts of perceptual learning, inhibit processes associated with attention weighting. In the discrimination condition, a similar pattern emerged where those in the simple image complexity condition had significantly lower overgeneralization and significantly higher correct exception responses than the other three complexity conditions. In conditions of increased complexity participants continued to overgeneralize roughly half of the exception items presented.

In addition to item response patterns in training, there were also hypothesized patterns in the transfer task that would support predictions of the ATRIUM model. Hypothesis 1c predicted that increased image complexity would result in a decrease in exception responses in transfer. While the data did not support the hypothesis, it is interesting to note that exception responses did increase in the transfer task less similar to the training image stimuli. In the extreme conditions (i.e., simple and full environment) there was a 32% increase in exception responses in

the transfer condition that used stimuli opposite of what they trained. The notion of using rule-based response strategies when presented with unfamiliar stimuli is not supported. Instead the findings support the idea that previous exposure plays a role in developing response strategies. This lends support for development of interpolated skill which relies on imprinting from previously exposed stimuli. In all, the concepts of perceptual learning closely followed behaviors predicted in the exemplar module within the ATRIUM model.

Hypothesis 1d focused more specifically on patterns of exception responses in the presence of either the primary or secondary cue, both of which were necessary in combination to warrant a correct exception response. If a respondent was focusing in on just one of the two cues in their responses, there would be a higher percentage of exception responses when that specific cue value was present in an image. The primary and secondary cue response patterns serve as a more specific measure of the perceptual learning concept of attention weighting, which closely mirrors the notion of dimension attention in the ATRIUM model, and provides a look at which cues were attended to in the transfer task. Although there were no differences between complexity training conditions, there were differences between the simple and full environment transfer tasks. The more complex transfer task (i.e., full environment transfer) resulted in more primary cue exception responses than secondary cue responses, while the pattern was flipped for the simple image transfer task (i.e., more secondary cue responses, less primary cue responses). The notion that responses will follow rule-based response patterns more when image complexity is increased was supported by this outcome.



Overall, the predictions of the model were largely supported. Two of the three criteria for model extension were supported. In training, the simple exposure training condition followed predictable patterns of responses. There were differences between the simple image complexity condition, and all of the increased image complexity conditions; however, the addition of more distracting visual cues did not magnify the reduced learning pattern further. In addition, there were some observable patterns of response in the transfer tasks that supported the use of rule-based response strategies with increased complexity. Overall, the findings support the notion that the theoretical model of category learning has value in predicting behavior as applied to complex real world task environments.

### *Hypothesis 2*

The results of Hypothesis 2 dealt with comparing the overall performance in training and transfer between training conditions. As alluded to in the beginning of this chapter, the results of Hypothesis 2 analysis were difficult to interpret. The subsequent patterns of results between the two training conditions appeared fundamentally different from the beginning. The discrimination training group consistently got more than 60% correct across all training blocks, but in transfer performed as poorly as a control group who received no training, and on average, only correctly identified 30% of the transfer items. For the exposure training group, there was a consistent improvement across training blocks. The transfer performance results showed that overall performance in the exposure group was essentially twice as good as both the discrimination and control groups. Should the two training groups have been eliciting learning of the same category learning skill, the logical pattern would be the condition that produced better

performance in training would have resulted in greater transfer task performance as well. Since the pattern did not hold, it is reasonable to presume that the training conditions were intrinsically different. The reason for the unexplainable performance differences could have resulted from a number of design and conceptualization factors, which will be further discussed in theoretical implications and study limitations. Overall the predictions of hypothesis 2 were not supported.

### *Hypothesis 3*

Due to the unexpected outcome of the analysis of differences in training type, the predicted interactions of Hypothesis 3 were obviously not supported. In spite of this, further analysis of the exposure training group revealed an interesting pattern of results. For the simple transfer task, performance was better when individuals trained on simpler training images. In fact, those in the full environment complexity training got 37% of the items correct, while those in the simple complexity training condition averaged 70% correct. In the full environment transfer, however, all complexity conditions performed nearly equally well. Taken further, breaking the training down into its simplest parts has training utility for complex task domains. At the same time the finding suggests individuals, who are trained in applied domains, have more difficulty identifying the underlying principles involved in the task.

### Theoretical Implications

Several interesting theoretical implications resulted from this study. Despite some unpredicted outcomes in the transfer task, the study provided critical information that answer

some questions and generate several new ones. These implications are presented in this section as they related to outcomes of tests model fit, training type, and transfer performance in the exposure training condition.

### *Model Fit*

The simple exposure condition which served as the bridge between the simple stimuli previously used to test the ATRIUM and the more complex image conditions, followed the expected pattern of early rule-based responses followed by more exception based responses. The results support ATRIUM model predictions which transfer beyond the box and line categorizations that had little real world value in the original studies. When distracting cues were added to increase complexity in training, responses followed the prediction that more rule-based response strategies would be used (i.e., higher overgeneralization, lower correct exceptions, more primary rule exception responses in transfer). What is interesting is that there was no statistical difference between the condition where only one distracting feature was visible and manipulated, and the condition where multiple distracting features were manipulated in an image with all environmental cues present. The model predictions began to falter as the number of distracters present was increased in the training imagery. There was little additive impact beyond the initial impact of including one distracting feature. Stated another way, learning to suppress one feature was equally as difficult as suppressing multiple distracting features in terms of the perceptual judgment task.

Another interesting outcome of the model fit analysis was the number of exception responses in transfer. There tended to be a pattern of higher number of exception responses on the transfer tasks that less resembled the images which individuals were trained on. On the surface, this appeared to have gone against the prediction of the model that rule-based responses would be more prevalent as complexity increased. The number of exception responses can be a deceiving measure though. In the conditions where individuals were more familiar with the stimuli they were tested with in the transfer task, they tended to have *less* exception responses. Taken further, category learning can be thought of as a process of obtaining balance between rule and exemplar responses to achieve a response strategy which optimizes categorization responses. If individuals are more familiar with stimuli, they are better equipped to correctly select between rule based and exception based categorizations. In all, the findings provide support for the importance of interpolative perceptual skill in category learning, since exception responses were less impacted by image complexity and more by familiarity with stimuli.

#### *Training Type*

The fact that the training type manipulation produced contradictory results, at first, appears to be a shortcoming of the design of the study. Actually this outcome touches on an important theoretical difference between perceptual and cognitive learning mentioned briefly earlier in the dissertation. The argument is whether or not perceptual and cognitive learning should be isolated in training. Evidence of the distinction manifests itself in terms of context. In order to control primed response bias, efforts were made to remove as much contextual information as possible from the training. As such, real world labels for the aviation task were

replaced with nonsense labels, strong visual cues in the real environment were altered (i.e., PAPI lighting), and the population consisted of almost entirely novice population with no prior aviation experience. The removal of context produced an environment where a real world task could be presented with low risk of conceptual priming. Theoretically, this produced an optimal environment for getting an unbiased picture of category learning, but a consequence of removing context is that the training conditions themselves served as the vehicle for generating context for the task. In the exposure group, lack of context was not an issue because the training task resembled the transfer task, and so the categorizations which they were performing in training provided direct transfer context for response in transfer. In the discrimination training, the same/different categorizations performed did not provide context for the transfer tasks. As evidenced by the higher overall performance scores in training, there is no question that individuals in the discrimination group were able to accurately distinguish between same and different images by category. Yet, the discrimination training condition was no better than a no training control in the transfer task. In essence, the discrimination condition could have elicited learning within the context of an entirely different category labeling scheme. If that were true, instead of learning in terms of the four categories introduced in pre-training, individuals in the discrimination training would be more adept at labeling in terms of two categories: same and different. Without training in the context of the four category labels, simply knowing whether two images were same or different would not be sufficient for categorization of the intended four category labels. In other word, rule-based response strategies developed in discrimination training could not readily translate into improved task performance without additional

information. What is unknown is whether a similar pattern of response would occur for an experienced group of pilots who are familiar with the task.

### *Transfer Performance Differences in Exposure Training*

Although the comparisons in Hypothesis 3 were not supported as well, the results observed across complexity conditions from the exposure training group provided some important findings. When tasked with a full environment transfer task, training complexity had no noticeable impact on performance. In the simple transfer task however, there were noticeable declines in performance as image complexity in training increased, or became more dissimilar from the transfer stimuli. Performers in the high complexity training conditions got less than half as many responses correct than those in the simple training condition. The implication is that providing training with reduced visual information has little impact on the ability of individuals to extrapolate critical cue information into more complex visual environments. On the other hand, individuals trained with more complex images have a difficult time extrapolating what they learned when the task is simplified into only its essential parts. Taken further, the ability to break down a task into its essential parts may be better trained using images that provide direct practice with only the critical environmental cues. In some ways, the results lend support that there are directional differences for perceptual skill development as complexity increases. It is easier for an individual, who is only given critical information, to extract information from a more complex visual scene than if the reverse is true. The influx of additional distracters in training detract from the ability to break down the task into its most basic parts, thus training the essentials of a perceptual task are better suited using training which focuses more on interpolated

skill. The result of this finding has important practical implications for training which are further discussed in the following section.

The theoretical implications from this study provide valuable information toward better understanding perceptual and category learning theory. The study provides support for predictions of the ATRIUM model when applied into a more complex real world context. In addition, the lack of transfer for the discrimination training condition provides an interesting illustration of the distinction between perceptual and cognitive learning and sheds light on the importance of context for novices in training. Finally, evidence for the additive value of task isolation in exposure training was observed.

### Practical Implications

Equally important to tying the dissertation to a theoretical base, are the practical implications associated with it. Most obviously, the goal of the study was to bridge the gap between a theory of category learning which is grounded in categorization tasks that have limited generalizability to the real world, and a training technique for a complex real world aviation task. There are a number of important practical implications which should be noted.

One of the underlying practical goals of the research was to explore cost effective methods of training individuals in complex tasks that require quick and accurate perceptual judgments. The visual approach aviation task used in the study is a good example of a complex task environment. The range of complex skills necessary for flight, in addition to the time and resources required to train these skills, cannot be addressed by simply adding more simulator or

class time to the training footprint. Due to the obvious limitations, developing cost and time efficient methods of skill training is important. Exploring methods that can be developed using commercial-off-the-shelf software which can be easily deployed to trainees, and still adequately train an underemphasized skill (e.g., perceptual skill) in training, can be extremely valuable tools. The training modules should be considered an abbreviated sample of the types of visual stimuli which would occur in a visual approach. In just about 30 minutes of training, novices in the exposure training group showed learning gains toward making perceptual judgments of a complex aviation task. The study serves to support the potential low cost training modules can provide in complex task environments.

Second, there was an interesting outcome resulting from the varied complexity conditions within the exposure training. One of the age old questions in complex task environments is what level of fidelity is required for optimal skill acquisition. When given a transfer task that included both relevant and irrelevant visual information, it did not matter if individuals had been trained with all of the information, or with only the information relevant to the categorization task. On the other hand, individuals who trained with higher complexity images were unable to accurately identify categories when distracting context was removed. One can deduce that as long as the task is accurately represented in training, novices do not need to train using images that depict the full range of irrelevant cues in the environment. Logically, training with task isolation would be useful in instances where task performance could occur in both visually complex environments, and where only degraded visual information is available. In the visual approach



aviation context, isolation training would be useful for nighttime approaches where limited visual cues are available. Based on the findings of this study, an individual trained to distinguish between glideslope variations using full environment images, will experience difficulty when asked to complete a similar task in commonly occurring scenarios (e.g., nighttime or desert, snow, and ocean terrain) when previously viewed cues are unavailable. Taking the example further, inappropriate training of perceptual judgment could lead to catastrophic outcomes.

The focus of the study was on method and theory as opposed to generalizing to developing aviation skill. Regardless of overall goals, it is encouraging that a novice population was able to improve performance on an unfamiliar aviation task with very little additional context. Comparatively low scores on the transfer task, further lends credence to the commonly expressed view in the aviation industry that the visual approach is difficult to train. Practically speaking, the study serves as evidence that individuals can be trained to make difficult perceptual judgments with easily deployable training methods.

#### Limitations and Future Research

Overall, a concerted effort was made to minimize study limitations through design, but there are a few limitations worth mentioning. Although the limitations were not determined to be severe enough to confound the results, they should still be taken into account when considering the generalizability of the study. Where applicable, suggestions for future research to address these limitations are made.

The first limitation of this study was that an exact replication of the ATRIUM model task was not used. Including an ATRIUM task condition would have provided a better baseline for comparing the simple exposure group to the original task. Including a baseline would have strengthened the assertion that the simple exposure condition served as an acceptable replication of the original task in a real world application. The resulting response patterns of the simple exposure condition followed model predictions close enough that the lack of the original task condition was determined to be of minimal overall impact.

A second limiting aspect of the study to consider was the distracting cue manipulation in the study. An effort was made to create a step-like progression of image complexity from the simplest image to the full environment images to get a gauge for the additive impact of complexity on the task. Since complexity is difficult to quantify into quantitatively incremental increases, determining equal distance between conditions is challenging. Despite the psychometric limitation, a logical conception of complexity increase as related to the aviation task domain was utilized in the study. There are countless visual cues in the environment which could have been manipulated, that have been shown to impact visual skill. By limiting the manipulation of the visual scene to three cues (i.e., runway size, airport size, and overall terrain density), a qualitative instead of quantitative approach to defining complexity increase resulted in minimal limiting factors of the complexity condition. The results seemed to suggest there was no additive impact of increasing complexity beyond the introduction of one distracting feature

into training. The lack of difference between higher complexity conditions further alleviated concern that uneven distance between cues had an uneven impact on the overall results.

A third limitation of the research was lack of adequate context for the discrimination training condition. In an effort to obtain a “pure” measure of category learning, effort was made to remove contextual information which might prime participant responses. As discussed previously, the lack of context was thought to have contributed to the exceptionally low transfer scores in the discrimination training group. The exposure training condition essentially received practice on the transfer task through training. Since the discrimination group did not receive practice and had limited context to reference, the transfer task was too difficult for novice participants. As a result of the imbalance in experience with the transfer task between training groups it is difficult to draw conclusions on differences between groups.

Before drawing definitive conclusions on the utility of the exposure and discrimination training techniques, a follow up study which includes more contextual cues should be conducted. The discrimination training method has been shown to enhance perceptual skill in other domains. One method of further testing the effectiveness of discrimination training would be to use an experienced pilot population with the same task. The current study used a student population with virtually no aviation experience. By testing an experienced pilot population with the same task, however, previous experience could provide additional context which would impact the ability to transfer skills learned using the discrimination task into the transfer task.

A fourth limitation of the research deals with the temporal aspects of the development of perceptual skill. Previous research suggests there may be a temporal lag in the development of extrapolated (i.e., rule-based) perceptual skill. Since the administration of the study took place in one sitting, the possibility remains that measures may not have accurately captured extrapolated skill development. Future research utilizing an immediate transfer and 24 hour time delay transfer task would help to identify whether there would be a shift in rule-based responses.

Finally, it is worth noting that the task was intended to replicate a real world complex task environment. Although the Microsoft Flight Simulator X software used to develop the software is considered a high visual fidelity simulation of aviation, no calibrations of display size and viewing distance were conducted to insure 1 to 1 visual replication of the distances involved in the real world task. This was mainly due to the limited resource of available space and laptops to administer the study. Additionally, the visual approach which is a maneuver which occurs in a highly dynamic aviation environment was represented using static imagery. Although a dynamic transfer task may have been even more informative to the transfer to the aviation task, resources for stimuli development, and complexity involved with developing an adequate number of accurate dynamic transfer task items, made it impractical to address this issue. Knowing in advance that the majority of the population participating in the study would not have prior aviation experience, the limitations were determined to be acceptable concessions of task realism.

Despite the inability to produce a dynamic transfer task for this study, it would be worthwhile to consider development of a dynamic performance measure of visual approach for future investigations. In many perceptual learning domains it is not uncommon to train via static image exposure, but in a domain which requires perception of a continuously changing environment, it is difficult to say whether training with a series of snapshots provide perceptual skill which readily transfers to the dynamic environment.

### Conclusion

The goal of the study was to investigate how well a model of category learning predicts response behavior in applied contexts, and to test the effectiveness of a specific type of perceptual training. While not all of the hypotheses put forth in this paper were supported, the findings resulted in valuable insight for training, perceptual judgment, and category learning. More specifically, response patterns in a more complex perceptual judgment task largely supported predictions of a theoretical model of category learning. Also, lack of transfer of the discrimination training condition highlighted the importance that context plays in perceptual judgment tasks. Additionally, image complexity in training led to differences in transfer that provided support for task isolation training. In all, the findings of this study contributed to both theoretical and applied aspects of perceptual learning and helped bridge some of the gap that exists between them.



## **APPENDIX A: IRB APPROVAL LETTER**



University of Central Florida Institutional Review Board  
Office of Research & Commercialization  
12201 Research Parkway, Suite 501  
Orlando, Florida 32826-3246  
Telephone: 407-823-2901 or 407-882-2276  
[www.research.ucf.edu/compliance/irb.html](http://www.research.ucf.edu/compliance/irb.html)

### Approval of Human Research

From: UCF Institutional Review Board #1  
FWA00000351, IRB00001138  
To: Michael T. Curtis  
Date: February 16, 2011

Dear Researcher:

On 2/16/2011, the IRB approved the following human participant research until 2/15/2012 inclusive:

Type of Review: UCF Initial Review Submission Form  
Project Title: Perceptual judgment: The role of image complexity and training method on category learning  
Investigator: Michael T. Curtis  
IRB Number: SBE-11-07465  
Funding Agency: Federal Aviation Administration  
Grant Title: Broadening the scope of AQP through training evaluation and development  
Research ID: N/A

The Continuing Review Application must be submitted 30 days prior to the expiration date for studies that were previously expedited, and 60 days prior to the expiration date for research that was previously reviewed at a convened meeting. Do not make changes to the study (i.e., protocol, methodology, consent form, personnel, site, etc.) before obtaining IRB approval. A Modification Form cannot be used to extend the approval period of a study. All forms may be completed and submitted online at <https://iris.research.ucf.edu>.

If continuing review approval is not granted before the expiration date of 2/15/2012, approval of this research expires on that date. When you have completed your research, please submit a Study Closure request in IRIS so that IRB records will be accurate.

Use of the approved, stamped consent document(s) is required. The new form supersedes all previous versions, which are now invalid for further use. Only approved investigators (or other approved key study personnel) may solicit consent for research participation. Participants or their representatives must receive a copy of the consent form(s).

In the conduct of this research, you are responsible to follow the requirements of the Investigator Manual.

On behalf of Joseph Bielitzki, DVM, UCF IRB Chair, this letter is signed by:

Signature applied by Joanne Muratori on 02/16/2011 04:09:02 PM EST

IRB Coordinator



## **APPENDIX B: BIOGRAPHICAL DATA QUESTIONNAIRE**

1. Gender: Male \_\_\_\_\_ Female \_\_\_\_\_

2. Age: \_\_\_\_\_

3. Visual Acuity: \_\_\_\_\_

4. Do you use prescription corrective lenses (glasses or contact lenses)? Yes \_\_\_\_\_ No \_\_\_\_\_

(if you answer no please skip to question 5, if yes proceed to questions 4a – 4c)

4a. What type of vision do your lenses correct for?

Near-sighted \_\_\_\_\_ Far-sighted \_\_\_\_\_ Both \_\_\_\_\_

4b. Are you wearing your corrective lenses now? Yes \_\_\_\_\_ No \_\_\_\_\_

4c. If no, please explain: \_\_\_\_\_

5. Have you ever operated an aircraft? Yes \_\_\_\_\_ No \_\_\_\_\_

(if you answer no please skip to question 6, if yes proceed to questions 5a – 5d)

5a. How many total flight hours do you have? \_\_\_\_\_

5b. Do you have a pilots license? Yes \_\_\_\_\_ No \_\_\_\_\_

5c. How many years have you had your license? \_\_\_\_\_

5d. What type of license do you have? \_\_\_\_\_

6. On average how many hours per week do you spend playing video games?

0 \_\_\_\_\_ 1-4 \_\_\_\_\_ 5-9 \_\_\_\_\_ 10+ \_\_\_\_\_

7. Have you ever used Microsoft Flight Simulator? Yes \_\_\_\_\_ No \_\_\_\_\_

-Please rate your experience with Microsoft Flight Simulator:

No Experience					High Experience
1	2	3	4	5	

## **APPENDIX C: SPATIAL ORIENTATION MEASURE**

### INSTRUCTIONS

This is a test of your ability to see changes in direction and position. In each item you are to note how the position of the boat has changed in the second picture from the original position in the first picture.

Here is Sample Item 1.

These bars represent the boat's prow.

This is the correct answer. It shows that the prow of the boat has dropped below the aiming point.

(If the prow had risen, instead of dropped, the correct answer would have been C, instead of D.)

These are the five possible answers to the item.

This is the prow (front end) of a motor boat in which you are riding.

This is the aiming point. It is the exact spot you would see on land if you sighted right over the point of the prow.

This is the same aiming point shown above. Note that the prow has dropped below it.

Sample Item 1

To work each item: **First**, look at the top picture and see where the motor boat is headed. **Second**, look at the bottom picture and note the CHANGE in the boat's heading. **Third**, select the answer that shows the same change.

Try Sample Item 2.

This also shows that the prow of the boat is to the right of the aiming point. So, it is the correct answer.

(If the boat had turned to the left, instead of to the right, the correct answer would have been A.)

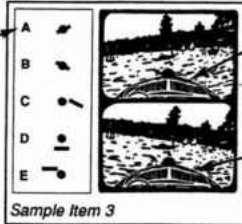
This is the aiming point.

This is the same aiming point. The motor boat is now headed to the right of it.

Sample Item 2

Now try Sample Item 3

This is the correct answer. It shows that the motor boat changed its slant to the left, but is still heading toward the aiming point.



Here the motor boat is slanted slightly to the right. (Note that the horizon appears to slant in the opposite direction.)

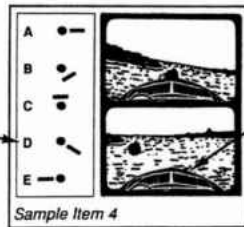
Here the boat has changed its slant toward the left. (To become level, the boat slanted back toward the right.)



Imagine that these pictures were taken with a motion picture camera. The camera is fastened rigidly to the boat so that it bobs up and down and turns and slants with the boat. Thus, when the boat tips or slants to the left (as in the lower sample in SAMPLE ITEM 3), the scene through the camera view finder looks slanted like this.

Look at Sample Item 4.

D is the correct answer. It shows that the boat changed its heading both downward and to the right; also that it changed its slant toward the right.

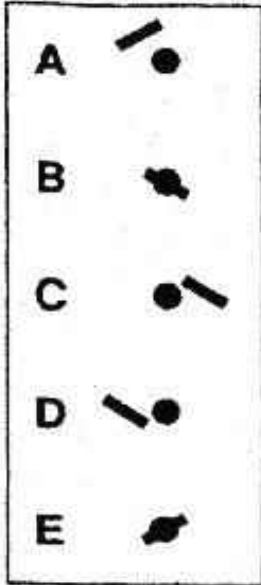
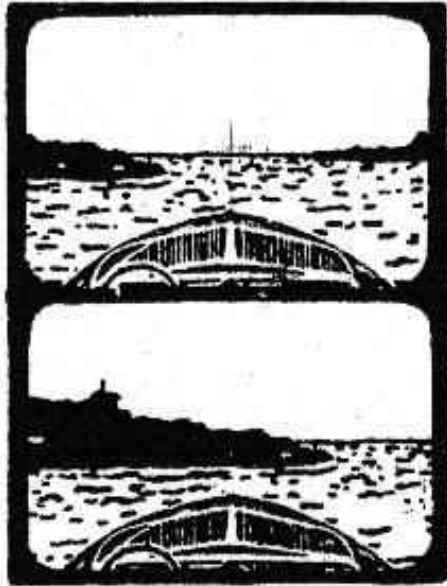
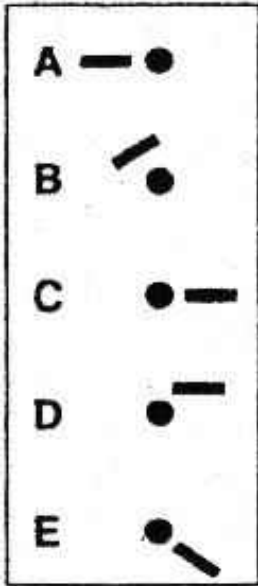


The prow of the boat has moved downward and toward the right. Also, it has changed its slant toward the right.

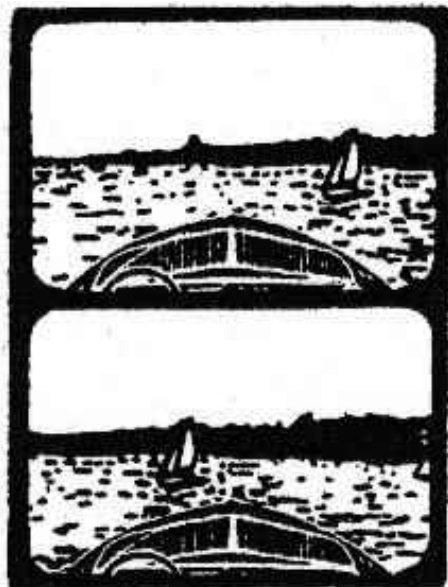
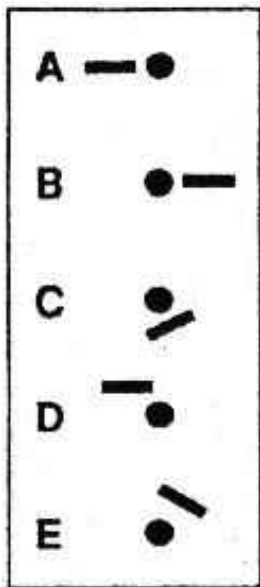
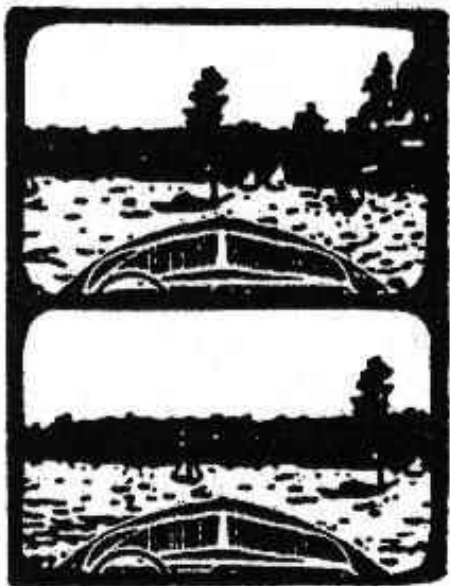
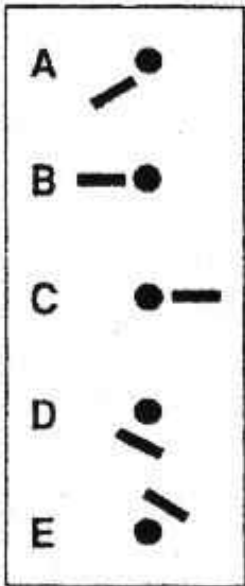
<p>A </p> <p>B </p> <p>C </p> <p>D </p> <p>E </p>		<p>A </p> <p>B </p> <p>C </p> <p>D </p> <p>E </p>		<p>A </p> <p>B </p> <p>C </p> <p>D </p> <p>E </p>				
<p>Item 5</p>	<p>Item 6</p>	<p>Item 7</p>						
<p>C is the correct answer. The prow appears to have moved to the left and downward. It has not changed its slant.</p>			<p>B is the correct answer. The prow appears to have moved to the left and downward. Also, it has changed its slant to the left.</p>			<p>E is the correct answer. The prow appears to have moved upward, and to have tipped left. It has not turned.</p>		

Sample of Items:

<p>A </p> <p>B </p> <p>C </p> <p>D </p> <p>E </p>	
---	--







## **APPENDIX D: TRAINING PERFORMANCE PLOTS**

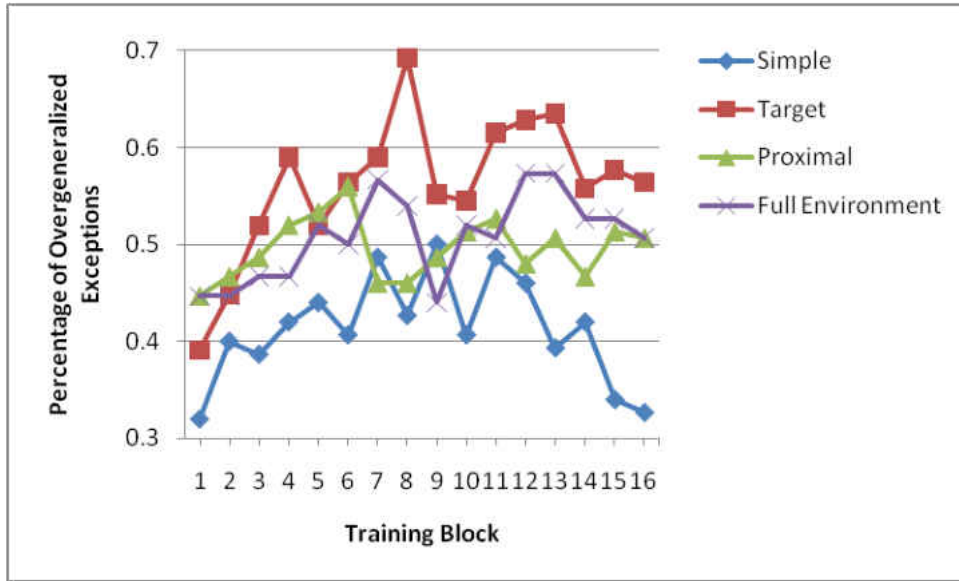


Figure 13. Percentage of overgeneralized responses across training blocks by training image complexity condition for the exposure training condition.

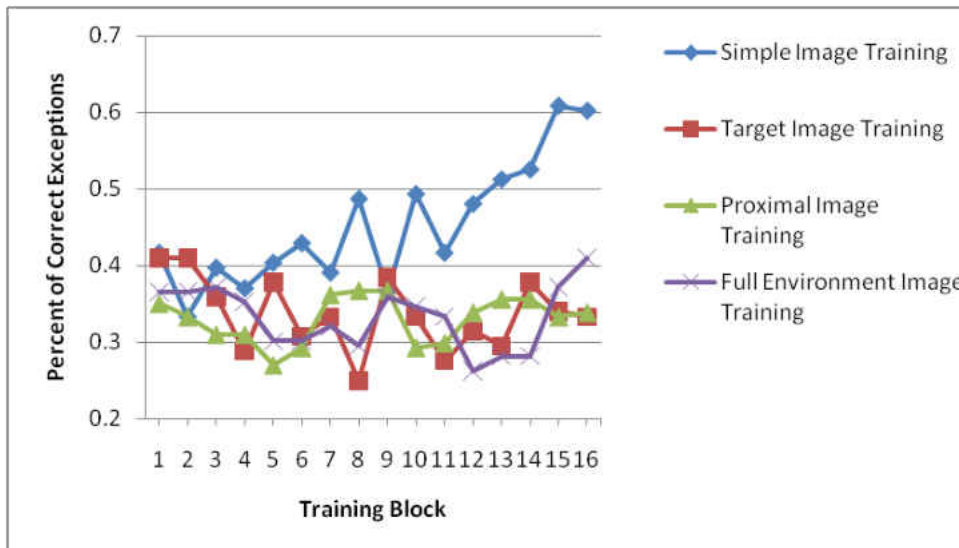


Figure 14. Percentage of correct exception responses across training block by image complexity training condition in exposure training condition.

## LIST OF REFERENCES

- Ahissar, M., Nahum, M., Nelken, I., & Hochstein, S. (2009). Reverse hierarchies and sensory learning. *Philosophical Transactions of the Royal Society B* , 364, 285-299.
- Anderson, J. R., Fincham, J. M., & Douglass, S. (1997). The roles of examples and rules in the acquisition of a cognitive skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition* , 23 (4), 932-945.
- Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology* , 56, 149-78.
- Biederman, I., & Shiffrar, M. M. (1987). Sexing day-old chicks: A case study and expert systems analysis of a difficult perceptual-learning task. *Journal of Experimental Psychology: Human Perception and Performance* , 13, 640-645.
- Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology* , 4, 55-81.
- Cooper, L. A., & Podgorny, P. (1976). Mental transformations and visual comparison processes: Effects of complexity and similarity. *Journal of Experimental Psychology: Human Perception and Performance* , 2 (4), 503-514.
- Curtis, M. T., Jentsch, F. G., & Maraj, C. (2009). How differences in spatial ability influence inexperience users in a visual perceptual aviation task. *15th International Symposium for Aviation Psychology*. Dayton.
- Czerwinski, M., Lightfoot, N., & Shiffrin, R. M. (1992). Automatization and training in visual search. *The American Journal of Psychology* , 105 (2), 271-315.
- Darby, R. (2006). Commercial jet hull losses, fatalities rose sharply in 2005. *Aviation Safety World* , 1 (2), 51-53.
- Doane, S. M., Alderton, D. L., Sohn, Y. W., & Pellegrino, J. W. (1996). Acquisition and transfer of skilled performance: Are visual discrimination skills stimulus specific? *Journal of Experimental Psychology: Human Perception and Performance* , 22 (5), 1218-1248.
- Doane, S. M., Sohn, Y. W., & Schreiber, B. (1999). The role of processing strategies in the acquisition and transfer of a cognitive skill. *Journal of Experimental Psychology: Human Perception and Performance* , 25 (5), 1390-1410.
- Drever, J. (1960). Perceptual learning. *Annual Reviews in Psychology* , 11, 131-160.
- Erickson, M. A., & Kruschke, J. K. (2001). Multiple representations in inductive category learning: Evidence of stimulus- and time-dependent representation. *Submitted for publication* .

- Erickson, M. A., & Kruschke, J. K. (2002). Rule-based extrapolation in perceptual categorization. *Psychonomic Bulletin & Review* , 9 (1), 160-168.
- Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General* , 127 (2), 107-140.
- Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review* , 100 (3), 363-406.
- Fahle, M., & Edelman, S. (1993). Long-term learning in vernier acuity: Effects of stimulus orientation, range and of feedback. *Vision Research* , 33 (3), 397-412.
- Fiore, S. M., Scielzo, S., Jentsch, F., & Howard, M. (2006). Understanding performance and cognitive efficiency when training for x-ray security screening. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting* (pp. 2610-2614). Santa Monica, CA: HFES.
- Fisher, D. L., & Tanner, N. S. (1992). Optimal symbol set selection: A semiautomated procedure. *Human Factors* , 34 (1), 79-95.
- Fisher, D. L., & Young, C. (1987). Perceptual learning: Theory and practice. *Proceedings of 31st Annual Meeting of the Human Factors and Ergonomics Society* (pp. 1049-1053). Santa Monica, CA: HFES.
- Gauthier, I., & Tarr, M. J. (1997). Becoming a "greeble" expert: Exploring mechanisms for face recognition. *Vision Research* , 37 (12), 1673-1682.
- Gibson, E. J. (1963). Perceptual Learning. *Annual Reviews in Psychology* , 14, 29-56.
- Gibson, J. J., & Gibson, E. J. (1955). Perceptual learning: Differentiation or enrichment? *Psychological Review* , 62 (1), 32-41.
- Goldstone, R. L. (2003). Learning to perceive while perceiving to learn. In R. Kimchi, M. Behrmann, & C. Olson (Eds.), *Perceptual organization in vision: Behavioral and neural perspectives*. (pp. 233-278). New Jersey: Lawrence Erlbaum Associates.
- Goldstone, R. L. (1998). Perceptual Learning. *Annual Reviews in Psychology* , 49, 585-612.
- Goldstone, R. L., & Barsalou, L. W. (1998). Reuniting perception and conception. *Cognition* , 65, 231-262.
- Gruson, L. M. (1988). Rehearsal skill and musical competence: Does practice make perfect? In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 91-112). Oxford, England: Clarendon Press.

- Guilford, J. P., & Zimmerman, W. S. (1948). The Guilford-Zimmerman aptitude survey. *Journal of Applied Psychology* , 32, 24-34.
- Haider, H., & Frensch, P. A. (1996). The role of information reduction in skill acquisition. *Cognitive Psychology* , 30, 304-347.
- Hoffman, R. R., & Fiore, S. M. (2007, May/June). Perceptual relearning: A leverage point for human-centered computing. *IEEE Intelligent Systems* , 79-83.
- Holzman, P. S., & Klein, G. S. (1954). Cognitive system-principles of leveling and sharpening: Individual differences in assimilation effects in visual time-error. *Journal of Psychology* , 37, 105-122.
- Jacobs, R. A., Jordan, M. I., Nowlan, S. J., & Hinton, G. E. (1991). Adaptive mixtures of local experts. *Neural Computation* , 3, 79-87.
- Jacobs, R. A., Jordan, M. I., Nowlan, S. J., & Hinton, G. E. (1991). Adaptive mixtures of local experts. *Neural Computation* , 3, 79-87.
- Kalinowski, A. G. (1985). The development of Olympic swimmers. In B. S. Bloom (Ed.), *Developing talent in young people* (pp. 139-192). New York: Ballantine Books.
- Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Science*, 88, pp. 4966-4970. USA.
- Kerr, R., & Booth, B. (1978). Specific and varied practice of motor skill. *Perceptual Motor Skills* , 46 (2), 395-401.
- Klein, G. A., & Hoffman, R. R. (1992). Seeing the invisible: Perceptual-cognitive aspects of expertise. In M. Rabinowitz (Ed.), *Cognitive science foundations of instruction* (pp. 203-226). Mahwah, NJ: Erlbaum.
- Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review* , 99 (1), 22-44.
- Kruschke, J. K. (2008). Models of categorization. In R. Sun (Ed.), *The Cambridge Handbook of Computational Psychology* (pp. 267-301). New York: Cambridge University Press.
- Livingston, K. R., & Andrews, J. K. (1995). On the interaction of prior knowledge and stimulus structure in category learning. *The Quarterly Journal of Experimental Psychology* , 48A (1), 208-236.

- Livingston, K. R., Andrews, J. K., & Harnad, S. (1998). Categorical perception effects induced by category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* , 24 (3), 732-753.
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review* , 95 (4), 492-527.
- Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review* , 111 (2), 309-332.
- Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review* , 85 (3), 207-238.
- Melcher, J. M., & Schooler, J. W. (1996). The misremembrance of wines past: Verbal and perceptual expertise differentially mediate verbal overshadowing of taste memory. *Journal of Memory and Language* , 35 (2), 231-245.
- Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General* , 115 (1), 39-57.
- Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review* , 101 (1), 53-79.
- Oakes, W. F. (1955). An experimental study of pitch naming and pitch discrimination reaction. *Journal of Genetic Psychology* , 86, 237-259.
- Op de Beeck, H., Wagemans, J., & Vogels, R. (2003). The effect of category learning on the representation of shape: Dimensions can be biased but not differentiated. *Journal of Experimental Psychology: General* , 132 (4), 491-511.
- Palmeri, T. J. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition* . , 23, 324-354.
- Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology* , 77 (3), 353-363.
- Scardamalia, M., & Bereiter, C. (1991). Literate expertise. In K. A. Ericsson, & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 172-194). Cambridge, England: Cambridge University Press.
- Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science* , 3 (4), 207-217.
- Schneider, W., & Detweiler, M. (1988). A connectionist/control architecture for working memory. *Psychology of Learning and Motivation* , 21, 53-119.

- Schyns, P. G., & Rodet, L. (1997). Categorization creates functional features. *Journal of Experimental Psychology: Learning, Memory, and Cognition* , 23, 681-696.
- Shappell, S. A., & Wiegmann, D. A. (2003). *A human error analysis of general aviation controlled flight into terrain accidents occurring between 1990-1998*. Springfield, VA: National Technical Information Service.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review* , 84 (2), 127-190.
- Smith, E. E., & Medin, D. L. (1981). *Categories and Concepts*. Cambridge, MA: Harvard University Press.
- Tajfel, H. (1957). Value and the perceptual judgment of magnitude. *Psychological Review* , 64 (3), 192-204.
- Wilson, D. A., & Stevenson, R. J. (2003). Olfactory perceptual learning: The critical role of memory in odor discrimination. *Neuroscience & Biobehavioral Reviews* , 27 (4), 307-328.
- Wulf, S. (1938). Tendencies and figural variations. In W. Ellis (Ed.), *A Source Book of Gestalt Psychology*. New York, New York: Harcourt Brace and World.