NORTHWESTERN UNIVERSITY

Differences in Conceptual Organization Among Types of Wine Experts:

The Impact of Goals on Representation

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

For the degree

DOCTOR OF PHILOSOPHY

Field of Psychology

By

Julia Beth Proffitt

EVANSTON, ILLINOIS

December 2008

ABSTRACT

Differences in Conceptual Organization Among Types of Wine Experts:

The Impact of Goals on Representation

Julia Beth Proffitt

Most cognitive research on conceptual structure has studied undergraduate populations and either natural (biological) or artificial (experiment-specific) categories. This project investigates how people with extensive, rich knowledge about a complex real-world domain organize and use that knowledge. The research extends prior work on differences among types of experts within biological domains (e.g., Medin et al., 1997, Proffitt et al., 2000, Medin et al., 2005) to explore these issues within a non-biological domain (wine). Because objects in this domain do not fall into a clear hierarchical taxonomy, this research provides an opportunity to explore issues related to the cross-classification of objects (Ross & Murphy, 1999). Another major focus of the research is the impact of differences in goals and domain-related activity on conceptual organization and use.

The researcher interviewed 3 types of wine experts (connoisseurs, retailers, and winemakers) to assess their expertise and identify behavioral differences among the groups. The 30 experts sorted 40 wine labels and generated category names for each group they created in a multi-level hierarchy and had the opportunity to repeat the process up to 2 more times. To test the robustness of the categories generated during the first session, several months later 24 of the

original experts completed a series of tasks including category membership and typicality judgments, similarity judgments, and an open-ended inference task.

In their sorts, the most common category types focused on color, region, and grape varietal. A common pattern was the combination of multiple category types—within a sort, within a level of a sort, and even within a category. The results of the second session found that the leading categories from the sorts were also important for the inference tasks, but the type of property, more than the type of expert, influenced which types of categories drove the inference.

ACKNOWLEDGMENTS

I owe my completion of this project to a large group of people. Most important, of course, is my advisor, Doug Medin, who so generously gave me an academic home at Northwestern and has exhibited extraordinary patience and support during my extended career as a graduate student. His wisdom and guidance have been invaluable and long will be appreciated. Committee members Lance Rips and Satoru Suzuki have also been incredibly generous with their time and tolerant of my unusual needs. I feel privileged to have worked with three people of such integrity and intelligence.

My sincere appreciation goes to the faculty and staff of the BOS Department of Franklin & Marshall College (and the institution as a whole). As I worked on this research almost entirely at a distance from my home institution, the support of my colleagues here—both emotionally and logistically—truly facilitated my work.

Thank you to Mark Chien, William Gulvin, Carl Helrich, Bradley Knapp, Lenny Koch, Joanne Levengood, Stephen Menke, Eric Miller, Marnie Old, Stephen Olin, Robert Peters, Edward Potter, Patrick Raynal, Edward Sands, Joe Spurlock, Amy Thorn, Eileen Tobias, all of the other wine experts and novices who so generously gave their time and shared their expertise with me.

There are so many other people who have helped me along the way: people like my mother, Judith Hackman, Gaye Wilson, Christine Donis-Keller, Raquel Klibanoff, and other family, friends and colleagues near and far who provided encouragement, companionship and moral support. A special thank you is due to Carol Ford for sharing her valuable talents at a much-needed time.

I could not have completed this work without the help and support of my family. My father has been a savior to me, providing pretty much every kind of support I needed, when I needed it. My mother-in-law, Elizabeth Jameson, gave up her life for months at a time to help with the children so I could work on this dissertation. And finally, to my wonderful husband Trex and our children, Catherine, Lauren, and Edward: thank you for everything you do, every day, to make this worthwhile. My deepest apologies for the hardship and neglect you have had to endure while I completed this project. I vow that someday (hopefully soon) I will arrange a play date, cook a meal, and maybe even take a shower.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE: INTRODUCTION

How do people use categories to organize their knowledge about the world? The question of how people judge entities to be the same kind of thing is a fascinating and important puzzle. However, classification is just the beginning. After deciding that something is a member of a particular category, a person can use that information to talk about it, think about it, make predictions about it, and do something with it. Although these category functions have been neglected relative to the problem of classification, there is a growing body of research on the relationship between conceptual organization and category use.

I predict that within a domain different types of expertise yield different category representations, which in turn affect the way domain knowledge is used in reasoning. In this dissertation, I conduct a set of experiments to investigate the conceptual organization and reasoning of three kinds of wine experts.

Research Questions

This research seeks to address four issues. First, how do people with extensive, rich knowledge about a complex real-world domain organize that knowledge? Most work on category structure has studied undergraduate populations and artificial (experiment-specific) categories. Similarly, research on the influence of category use on representation has emphasized short term, experimenter-defined goals and contexts. This project investigates how people with extensive, rich knowledge about a complex real-world domain organize and use that knowledge.

Second, do different kinds of experts within the same domain organize it differently? Although expertise research has emphasized differences among experts based on their level of expertise (novice-intermediate-expert-master) there is evidence of qualitative differences among experts, based on functionally defined groups (Medin, Lynch, Coley, & Atran, 1997; Medin, Ross, Atran, Cox, Coley, Proffitt, & Blok, 2005; Proffitt, Coley, & Medin, 2000). People who share substantially the same knowledge base (i.e., who could answer the same factual questions) but who differ in their orientation to the domain (i.e., their goals and the ways they interact with the objects in the domain) may differ in their representation and use of that knowledge. For example, although sommeliers and winemakers might both know a lot about a given wine, they may emphasize different aspects and as a result view the same wine as a member of different "families." Studies of tree and fish experts support the hypothesis that different ways of engaging with a domain can lead to very different conceptions of it.

Third, are multiple organizations of the domain available to individual experts? Models of concepts and reasoning frequently assume people have and use a single, hierarchical taxonomy to organize their knowledge. Although it may seem intuitively obvious that this is not the case, research documenting cross-classification is scarce.

Finally, are there differences in how experts with different goals in the domain use their knowledge to make predictions about unfamiliar examples and/or properties? In other words, are there differences in how they extend their knowledge? Assuming that multiple classification systems do emerge, how do individuals navigate these options when reasoning about the domain?

Empirical Context

This work explores how three types of wine experts organize their knowledge about wine and how they use it to make inferences. Wine lies at an interesting intersection of the natural and man-made world, in that it is a man-made product crafted from natural materials. Three features of this domain make it particularly appropriate for my investigation of goal-directed, category-based reasoning.

First, no single taxonomic structure naturally dominates the domain of wine. Instead, there are a variety of valid features that could be used to organize a sample of wines. For example, type of grape, vintage, terroir,[1] and wine maker all are relevant to decisions about wine quality. Wine is a complex domain in which cross-classification probably occurs.

Second, there are functionally distinct types of experts in the domain. I focus on three groups: connoisseurs, retailers, and winemakers. Connoisseurs drink and may collect wine but are not professionally involved in the wine trade. Retailers are professionals who sell wine. Winemakers are professionals who grow grapes and produce wine. While all three groups know a great deal about wine, their goals and patterns of engagement with the domain are likely to be different.

Third, the task of making inferences has strong ecological validity, as wine experts regularly (and routinely) make consequential decisions based on incomplete information. For example, connoisseurs must decide whether a bottle of wine they have never tasted is worth the price. Retailers must decide whether a particular wine is likely to please a certain type of customer. Winemakers must decide what procedures will enable an immature wine to develop its full potential.

In sum, the wine domain appears to be uniquely appropriate for addressing the research questions to be explored in this dissertation. Wines can be classified in a variety of ways. There

[1] Terroir refers to the specific geographic and climatic characteristics of the land on which the grapes were grown.

are diverse types of wine experts with different domain-related goals. Extending knowledge through inference is a realistic and natural activity for this domain.

Conceptual and Empirical Background

This dissertation concerns how wine experts organize and use their specialized knowledge about the domain. To set the stage for this study, I begin with a review of key issues in our understanding of concepts, focusing on the impact of category use and the phenomenon of cross-classification. I explore the relationship of these two topics to the expertise literature. This leads to the design of my study and the implications of existing research on wine expertise for this research.

Categories, Concepts, and Their Relations

What is a category? A common view is that categories refer to sets of the same kind of thing, thus "Categories are equivalence classes of different (i.e., discriminable) entities" (Sloutsky, 2003, p. 246). Medin and Rips acknowledge that a strict definition is elusive and offer this somewhat less rigid interpretation:

> Cognitive scientists generally agree that a concept is a mental representation that picks out a set of entities, or a category. That is, concepts *refer*, and what they refer to are categories. It is also commonly assumed that category membership is not arbitrary but rather a principled matter. (Medin & Rips, 2005, p. 1)

While the study of concepts and categories has often focused on physical object categories, there is widespread recognition that humans use a broader range of concepts, that may refer not only to objects, but also to people (e.g., Fiske, 1998), colors (e.g., Berlin & Kay, 1969), events (Rips &

Estin, 1998), or even patterns of dots (Posner & Keele, 1968), to name just a few possibilities (see Medin, Lynch, & Solomon, 2000, for a review).

The principles for organizing the varied "things" that make up categories are also varied, and our understanding of those principles has evolved over time (for reviews, see Murphy & Medin, 1985; Komatsu, 1992; Murphy, 2004; Medin & Rips, 2005). Early scholars subscribed to what is known as the classical view; according to this model, defining sets of necessary and sufficient features determine category membership. However, although people generally believe that definitional features exist for categories, they are usually hard-pressed to identify such features (Gelman, Coley, & Gottfried, 1994). Another problem for the classical view is evidence that category membership is graded—some examples are considered better members of the category than others (Smith, Shoben, & Rips, 1974). Rosch & Mervis (1975) observed that members of a category may not share any particular feature, but instead demonstrate a certain "family resemblance." These observations spurred the development of probabilistic views such as the prototype model in which potential category members are compared to a prototype—a feature-based representation of the central tendency of the category. Rosch and Mervis (1975) also noted that the more similar an instance is to all other category members, the more typical it is of the category.

These traditional models of concepts rely heavily on the evaluation of feature similarity, an emphasis that has extended to theories of semantic memory (Collins & Quillian, 1969; Smith et al., 1974) and inductive reasoning (Osherson, Smith, Wilkie, Lopez, & Shafir, 1990; Rips, 1975; Sloman, 1993). Yet judgments about similarity are quite variable and context dependent (Medin, Goldstone, & Gentner, 1993). Although similarity-based models are still quite popular, a strong case has been made for models that give people's theories and expectations about the world a prominent role (Komatsu, 1992; Murphy & Medin, 1985; Neisser, 1987). Not only can theory-based models constrain features and provide context for similarity judgments, they also are better at accommodating knowledge about diverse kinds of relationships among categories and features, such as causal relations (e.g., Ahn & Kim, 2001; Rips, 2001).

Another limitation of similarity-based views is that their applicability may be restricted to taxonomic categories. A taxonomy is

> a hierarchical system in which concepts differentiated into different levels of specificity (e.g., animal, dog, collie) are related by class inclusion e.g., a collie is a dog, a dog is an animal, a collie is an animal). Such a structure provides a rich network that supports inferences about categorical properties at various levels of specificity. (Lin, 1997, p. 3)

Hierarchical structures have long been a component of models of cognition (e.g., Collins & Quillian, 1969) and are certainly useful structures for organizing large quantities of information. Theoretically, they are cognitively efficient, because property inheritance reduces the need for redundant encoding and storage of information (but see Sloman, 1998).

Rosch et al.'s influential work on the basic level (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976) assumes hierarchical structure. These researchers identified an intermediate level of category, the <u>basic</u> level, that is privileged both linguistically and cognitively. In a taxonomy of furniture, the basic level category <u>table</u> would be more specific than its superordinate, <u>furniture</u>, and more abstract than it subordinate, <u>coffee table</u>. A variety of indicators converge on the basic level as special. Basic level names are the first learned by children, as well as the first offered by adults to describe an object. Objects from the same basic level category are similar in terms of overall shape and tend to share a large number of features. One explanation for the special status of the basic level stems from its differentiation: these categories are easy to identify and cognitively economical. What you know to be true of one member of the category is likely to hold for all members of that same category, but probably does not hold for members of contrasting categories at the same level.

Much of this research assumes that concepts fall into stable, hierarchical organizations that allow conceptual structure to be analyzed in terms of similarity and class inclusion relationships. Certainly many types of categories do seem to fit into taxonomies. Natural categories such as plants and animals are obvious examples.

However, while taxonomies are certainly real, there are reasons to suspect that hierarchic taxonomies are not a universal underlying conceptual structure. A purely taxonomic understanding of categorization simply "does not capture the entire spectrum and the richness of conceptual representations" (Lin, 1997, p. 6). Recent years have seen increased attention to categories such as thematic categories (Lin, 1997; Lin & Murphy, 2001; Murphy, 2001) and *ad hoc* categories (Barsalou, 1983) that depend on principles other than similarity in delineating category membership.

Acknowledgment of thematic categories reflects the observation that in the world, "Objects are not found organized by category but rather are embedded in spatial, temporal, and causal contexts. Such relational structures as events and themes are a common way of organizing information to make sense of what we encounter" (Markman, 1989, p. 37). Yet for many years, thematic categories were considered an immature category form that adults shed, a developmental pre-cursor to taxonomic categories (e.g., Inhelder & Piaget, 1964). More recent experiments (Lin, 1997; Lin & Murphy, 2001; Murphy, 2001) demonstrate that adults also choose to use thematic categories under certain conditions. For example, when asked to choose whether a thematic or a taxonomic match "goes best with X to form a category" (e.g., should a litter box or a lion be paired with a cat), 62% of responses were thematic (Lin, 1997).

Like thematic categories, ad hoc categories (Barsalou, 1983) violate the taxonomic norm and are not based on class-inclusion relationships. Instead, ad hoc categories are organized around goals, an example being "things to take from one's home during a fire." Members of this category (e.g., children, pets, stereo, blanket, wallet, photo album) are not similar in terms of material, size, shape, or purpose and thus "violate the correlation structure of the environment" (p. 211). Neither are they "well established in memory," (p. 211) yet they are still perceived as categories and share some characteristics—such as graded structure—with taxonomic categories.

Another problem with taxonomies is their assumption of mutual exclusivity among categories. Social categories, for example, are notoriously resistant to hierarchical organization (Lingle, Altom, & Medin, 1984). People tend to be members of multiple, overlapping categories (Kunda & Thagard, 1996), in which superordinate-subordinate relationships distinctions are inappropriate. Are Asian women a subgroup of Asians or of women? There is no "right" answer to this question. People can have multiple category labels applied to them. Whether and which categories are activated may depend on a number of factors, including cognitive load, temporary information-processing goals, and personal motivation (Macrae & Bodenhausen, 2000).

This phenomenon is not limited to social categories: cross-classification exists for artifact categories, abstract categories, *ad hoc* categories, and even natural categories. For example, the spreading yew plant is a kind of evergreen and a kind of bush. This dual-category membership presents a challenge for inference. Is something that is true of a yew more likely to be true of a pine or a hydrangea? There isn't really a single answer—it may depend on whether you are interested in seed dispersal or growing habits, both of which may have a genetic basis. The important point is that most people can find more than one compelling organization of things in the world.

Whereas social psychologists have grappled with the problem of cross-classification (see Macrae & Bodenhausen, 2000), cognitive psychologists have investigated the issue less intensively, usually only alluding to the phenomenon (Barsalou, 1982; Medin, et al., 1997).

A notable exception is Ross and Murphy's (1999) systematic investigation of category representations of food—a domain they suspected might be subject to cross-classification. The first stage of their study identified food categories by having undergraduates (1) generate categories for a set of common foods, (2) evaluate the membership of foods in a subset of the categories generated, and (3) sort foods into categories. These first three tasks revealed both taxonomic categories (kinds of food such as meats and beverages) and script categories (situation-derived groups such as appetizers and lunch foods), providing evidence of cross-classification. However, they noted, "it does not follow from this finding that people activate both kinds of categories when actually thinking about specific foods" (Ross & Murphy, 1999, p. 518).

Therefore they proceeded to test the activation of taxonomic and script categories in two experiments: (1) similarity ratings and (2) primed, speeded category verification. The design of these tasks drew on Barsalou's (1982) experiments exploring the distinction between context-independent and context-dependent category information. Context-independent properties are always activated with the concept, whereas context-dependent properties are only activated in certain situations. For example, for basketball "round" would be context independent but "floats" would be context dependent. Context-independent information should be equally available under a variety of conditions, whereas the context-dependent information should be more easily accessed when presented in the correct context.

In a similarity judgment task, Barsalou's (1982) participants judged the similarity of pairs of objects, some from *ad hoc* categories (e.g., flashlight and rope, both things to take on a camping trip) and others from common categories (e.g., sofa and desk, both furniture). Half of the subjects saw the category name alongside the item pair as they made their similarity judgment. For common categories, the properties that make members similar should be relatively accessible (context independent), whereas for *ad hoc* categories, the shared properties may be less accessible (context dependent). Indeed, priming context (by providing the category label) yielded higher similarity ratings for *ad hoc* categories, but not for common categories. The same pattern of results was obtained for the speeded property verification task.

Adapting Barsalou's (1982) two measures of context dependence/independence, Ross and Murphy (1999) demonstrated that the food script categories were not driven solely by task demands. Both types were available to subjects in other contexts. However, the taxonomic categories were more accessible than the script-based ones, which were in turn more accessible than novel *ad hoc* categories. In other words, script categories were activated for food items, but not to the same extent as taxonomic categories.

Ross and Murphy's (1999) final two experiments found that both script and taxonomic categories were used for making inferences, but that they were used differentially, with taxonomic categories licensing stronger inferences for biochemical properties but script categories having more inferential power for situational properties. For example, when choosing between a taxonomic and a script match in response to the probe, “Suppose that an enzyme, metacascal, had been found in [target food] in the country Quain. What food is more likely to contain metacascal?” people were more likely to choose a taxonomic match. Thus, given a target of “cookie,” they would choose “biscuit” over “ice cream.” But, in response to the question, “Suppose that [target food] is eaten at the annual initiation ceremony in the country Quain. What food is more likely to be eaten at the ceremony?” the script match was chosen more often. For this situational inference, “ice cream” was chosen over “biscuit.”

In sum, although the focus of many models and much research has been on taxonomically organized categories of objects, a full understanding of concepts and categories will need to account for a broader range of phenomena. These will need to address both a wider range in terms of category content and in terms of the principles relating categories and category members.

Category Use

Categories and concepts serve a number of important cognitive functions. Foremost among these functions are classification, communication, and inference. When people encounter something new, classifying it as a member of a known category is a natural, automatic part of the process of trying to recognize and understand it. Categories enable meaningful communication with others about the world. Knowledge of category membership in turn supports inference, problem solving, and productive thinking through conceptual combination.

Impact of Use on Conceptual Structure

Although cognitive psychologists widely acknowledge the importance of a variety of conceptual functions, research has emphasized the learning of novel categories and classification of objects. Diagnostic properties—those that distinguish members of different categories—play an important role in classification. As a result, studies focusing on the classification process have yielded descriptions of conceptual structure that emphasize diagnostic properties. However, different ways of using a category may affect the perceived importance of different kinds of features.

Researchers have begun to examine exactly how category use affects category representation (Jee & Wiley, 2004; Love, 2005; Markman, Yamauchi, & Makin, 1997; Ross, 1997; see Markman & Ross, 2003, for a review). In an artificial category learning study, Markman and Ross concluded that subjects who learned novel categories through inference tasks emphasized prototypical features and within-category relationships in their category representations. In contrast, those who had learned categories through standard classification tasks emphasized diagnostic features and between-category relationships. In another study, symptoms that predicted the effectiveness of a treatment were viewed as more predictive of a disease, even though other symptoms were equally so (Ross, 1997).

Research with real-world categories parallels this work, suggesting that experience using classifications to make decisions about action may affect conceptual structure. Murphy and Wright (1984) asked participants to list features of different categories of emotionally disturbed children. Experienced psychologists listed a smaller proportion of distinctive traits than novices did. In other words, experts were more likely to list the same feature for more than one category.

These findings differ from other research showing that experts generally have more differentiated subordinate level categories (Johnson & Mervis, 1997; Palmer, Jones, Hennessy, Unze, & Pick, 1989; Tanaka & Taylor, 1991).

What could be the reason for Murphy and Wright's anomalous results? Although they did not report the specific common traits that experts offered, it is possible that they were features that were tied to common treatments. So, for example, while the diagnosis of an aggressive-impulsive or depressed-withdrawn child may be different, the information most important to their caregivers may be the fact that both types of children tend to be sad and angry.

A more extreme stance is that category use drives the formation of concepts (Schank, Collins, & Hunter, 1986). In anthropology, there has been a long-standing debate on this issue. Are the categories that cultures use determined by form or function?

> On the one side are those who argue for an intellectual basis of folk biological classification: the folk are seen as natural historians interested in understanding organic diversity for its own sake (e.g., Levi-Strauss 1966; Berlin, Breedlove, and Raven 1973). On the other side are those who argue for the utilitarian basis of human attention in the environment, picturing folk biologists as pragmatists interested in the natural world primarily as a means for satisfying human needs (e.g., Malinowski 1954; Hunn 1982). (Boster & Johnson, 1989, p. 866)

In other words, do the categories exist "in the world" to be discovered by humans? Or does the way that humans interact with the world drive the categorization?

I will revisit this issue of how people's interactions with the world may affect their perception of it, but first I will present more information about a particularly important function of categories—their role in inductive inference.

Category-based induction

Research on inductive inference has focused on the use of taxonomic, categorical information (Osherson et al., 1990; Rips, 1975; Sloman, 1993). These similarity-based accounts have been used to explain phenomena such as the use of typicality and diversity in reasoning.

The typicality phenomenon describes a reasoning pattern observed by Rips (1975) in which people are more willing to generalize a novel property from a typical member of a category than from an atypical one. For example, the argument, “Dogs have property X; therefore all mammals have property X” would generally be considered stronger than the argument “Whales have property X; therefore all mammals have property X.” One explanation for this phenomenon is that because dogs are a more typical example of the mammal category, they are similar to more mammals. This similarity means the premise about dogs offers greater “coverage” of the conclusion category, mammals, resulting in the perception that the first argument is stronger (Osherson, et al., 1990). Their account of the diversity phenomenon is that an argument with more diverse premises (e.g., dogs and bats) provides more coverage of the conclusion than would less diverse premises (e.g., dogs and cats), resulting in the perception of greater argument strength for the diverse argument.

A number of studies have found that these strategies are less prevalent outside of typical American undergraduate populations (e.g., Lopez, Atran, Coley, Medin, & Smith, 1997, Medin et al., 1997). For example, in a study of tree experts, Proffitt et al. (2000) did not find strong evidence of these phenomena, but instead observed heavy reliance of the experts on causal-ecological strategies. Even when the tree experts used similarity-based strategies, they took a somewhat different form, which Proffitt et al. called *local coverage*:

> Instead it appears that experts view disease as spreading within smaller taxonomic groups, such as families of plants. As a result they frequently base their reasoning on what we call *local coverage*, which roughly constitutes a form of coverage but is based on a subset of the conclusion category. (p. 813)

Experts expected other trees highly similar to the premise trees (generally trees of the same genus) to be susceptible to the same disease.

Models of category-based inference have emphasized the role of similarity, yet some research suggest that a purely similarity-based account of inference is incomplete (Medin, Coley, Storms, & Hayes, 2003; Proffitt et al., 2000; Shafto, Kemp, Baraff, Coley & Tenenbaum, 2005). Similarity has been an important factor in models of inference, but more is needed to provide a complete picture, given evidence of non-similarity-based reasoning. Further, cross-classification presents a particularly interesting problem for models of category-based induction, "because the presentation of an item may access multiple categories and it is not clear how the categories accessed will influence the inductions" (Ross & Murphy, 1999, p. 500).

The Impact of Expertise on Conceptual Representation

How do knowledge and experience affect use and representation of concepts? Expertise represents extensive knowledge about a domain and exceptional performance on a complex set of domain-related tasks. But do experts just know more and process things faster? Or are there qualitative differences that result from expertise?

A major theme in the expertise literature has been to understand what underlies superior expert performance (for reviews see Glaser & Bassok, 1989; Gobet, 1998).[2] Several cognitive

---

[2] Note that this is a distinct issue from what it takes to become an expert (Ericsson & Charness, 1994; Ericsson, Krampe, & Tesch-Romer, 1993).

factors appear to drive expert-level performance. Prominent among these are (1) the development of automaticity (Anderson, 1987), (2) perceptual learning (Goldstone, 1998) and (3) changes to knowledge structures (Chase & Simon, 1973; Ericsson & Staszewski, 1989; Glaser & Bassok, 1989; Klein, 1997).

My focus is on how the acquisition of expertise affects knowledge structures. One well-documented effect is the chunking of information. With experience, people develop a larger repertoire of meaningful chunks that they can use to efficiently encode, retrieve, and organize information about the domain, be it chess (Chase & Simon, 1973), medicine (Lesgold, Rubinson, Feltovich, Glaser, Klopfer, & Wang, 1988), or wine (Hughson, 2003).

In addition to chunking information efficiently, experts may detect and emphasize different themes than novices do. In feature-listing tasks, experts have been found to list more implicit (behavioral or inferred) features for animals than explicit (observable, physical) features (Gobbo & Chi, 1986; Johnson & Mervis, 1997). A frequently-cited effect of expertise is the finding that experts organize concepts using deep structural principles, whereas novices rely on surface similarity.

Chi, Feltovich, and Glaser (1981) contrasted how experts and novices sorted physics problems, and found that the experts' sorts were organized around principles of physics whereas the novices' sorts were based on more superficial features such as the physical props involved in the problem (e.g., whether a pulley or a slope was involved). In a conceptual replication, Hardiman et al. (1989) directed experts and novices to judge the similarity of physics problems, pitting surface features against deep structure in a match-to-sample task. Consistent with Chi's findings, novices were more likely than experts to put together two problems that shared surface similarities.

Note, however, that the instructions in both studies emphasized deep structure. For example, Hardiman, Dufresne, and Mestre (1989) asked subjects to "indicate which of the two comparison problems 'would be solved most similarly' to the model problem" (p. 628). So while a real difference exists between the two groups, it may not be a difference of strategy, focus, or attention, but rather of knowledge; the selection criteria were successful at separating those who knew how to solve the problems from those who did not. Indeed, Hardiman et al. (1989) tested participants' problem-solving ability, confirming that novices had difficulty solving the physics problems. Not surprisingly, there was a correlation between categorization score and problem-solving score. Therefore, novices probably did not organize problems on the basis of solution methods because they didn't know how to solve them. The fact that they used surface features does not mean they preferred that approach; they did not have access to deep structure.

Neither do we really know what organization of the problems came naturally to the experts. True, they were <u>able</u> to group by solution methods. However, the fact that they could follow the instructions does not establish their preferred organization. Interestingly, Chi et al. (1981) report that experts took an average of six minutes longer than the novices to complete the sorting task. Presumably, the solution-based organization of problems was not one that was immediately "perceived" by them (as is sometimes claimed, e.g., Schoenfeld & Hermann, 1982), but one that required some effort. Differences in knowledge between experts are surely not imaginary. But to claim that experts simply know more about their domain than novices do is tautological.

In more open-ended tasks, experts appear to have access to multiple strategies. Comparing expert fishermen and novices, Shafto and Coley (2003) found that both groups generated taxonomic categories in sorting tasks, but experts generated significantly more categories organized around behavior and environmental considerations. In a subsequent inference task, there was an interaction between expertise level and property type (whether the inference was about disease or a blank property). Specifically, novices seemed to treat questions about diseases and blank properties as equivalent, extending them to comparable numbers of fish. Expert performance mirrored novices' for blank properties, but for diseases, experts extended the disease property to more fish. Subsequent analyses led Shafto and Coley to conclude that while the experts used taxonomic information for blank properties, they were using different information—knowledge about food chains—to guide their inferences about disease. In contrast, novices used taxonomic information for all properties; unlike the experts, the novices were *unable* to deploy different kinds of knowledge for different problems. This probably was due to a simple lack of knowledge. Like Chi et al.'s (1981) novices who were unable to sort physics problems on the basis of solutions they did not know, Shafto and Coley's novices may not have used the food chain relationships simply because they did not know them.

Boster and Johnson (1989) point out that the acquisition of expertise does not just affect the amount of knowledge, but also the kind of knowledge, stating that, "While morphological information is available to anyone who looks at organisms, cultural knowledge of the utility of the organisms usually requires extensive experience" (p. 868). If the information easily accessible to novices (e.g., for fish, shape) is also of use to experts, then experts may acquire multiple organizational systems.

Similarly, Boster and Johnson (1989) asked fishermen (from four regions) and novices to sort line drawings of fish into a set of mutually exclusive groups at a single level, with no limit on the number of categories. The pair-wise distances (aggregated by group) were correlated with the other groups' aggregate distances, as well as with distances based on scientific taxonomy, and with distances based on answers to a set of questions focusing on similarities in use or behavior of the fish. The researchers concluded that whereas novices based their sorts on taxonomy (derived from morphological information visible in the drawings), experts were, "apparently judging the similarities among the fish about equally on the basis of form and function" (p. 872). These correlational findings were supported by multidimensional scaling analyses of the distances among fish as well as by examination of the kinds of labels used; experts used both morphological and functional labels.

An additional question was whether the two groups were equal in terms of their variation. Generally anthropological research has found that more knowledge leads to greater consensus among informants. However, because the novices tapped a single type of knowledge (morphological), whereas the experts also tapped knowledge about functional information, Boster and Johnson (1989) did not expect increased expertise to lead to increased consensus. The Cultural Consensus Model (CCM, Romney, Weller, & Batchelder, 1986) works as an indication of knowledge when a <u>single</u> answer exists. In the case of a strongly cross-classified domain, there may be <u>multiple</u> correct responses. If they are equally preferred, there may not be evidence of consensus. Indeed, Boster and Johnson found greater consensus among novices than among experts because the novices were using a single factor—morphology—as the basis for their groupings. Although experts also used morphology, they exhibited more variation:

> Different experts appear to use functional and morphological information to different degrees in making their sorting decision; they not only show more variation in their sorts of the fish, they also offer more varied justifications for their sorts. (p. 880-1)

The emphasis of research on expertise has been to document and understand changes along a continuum from novice to expert,[3] but relatively little research has examined differences among types of experts within the same domain.

Expertise is more than just superior domain knowledge. Ericsson and Lehman (1996), describe expertise as "maximal adaptation to task constraints." Thus, different kinds of expertise should lead to different kinds of adaptations. These adaptations may lead experts to converge in their representations of the domain if their goals and behaviors rely on the same features (or highly correlated ones), but may result in diverging representations if the focal features are not aligned, or if multiple representations are available and useful.

Types of Expertise

In addition to distinctions among different levels of expertise, experts within a domain may be compared in terms of the type of expertise they have (Medin et al., 1997; Proffitt et al., 2000). The kinds of features needed to make fine distinctions among categories may differ depending on the domain as well—birders attend to behavioral characteristics, dog experts to physical ones (Tanaka & Taylor, 1991). Even within a domain, we find that different kinds of tree experts impose different organizations on a sample of trees, and that these organizations reflect the different priorities of their jobs (Medin et al., 1997). It is important to evaluate expert performance with respect to the specific tasks that characterize their endeavors.

---

[3] In addition to this common dichotomous contrast (between experts and novices), more differentiated gradations have been studied (e.g., Dreyfus, 1997; Johnson & Mervis, 1997; Melcher & Schooler, 1996; Patel, Arocha, & Kaufman,1994; Solomon, 1997).

As discussed earlier, how you interact with something can affect your representation of it. The memory model of transfer-appropriate processing (Morris, Bransford, & Franks, 1977) also suggests that the specific task or form of engagement with information affects what gets remembered about it. If this is true of the kind of short-term engagement typically involved in memory experiments and category learning studies, then extended intense engagement is likely to influence long-term knowledge structures. (Of course, the opposite is also a possibility—perhaps task only influences representation for short-term exposures. It may be that extended engagement allows one to discard contextual information and only encode core aspects of concepts.)

Yet although there is evidence that goals affect category representations (Barsalou, 1983, 1991; Markman & Ross, 2003), the stability of these effects is unclear. If, as the evidence suggests, the way that people use categories affects their representations of those categories, then it is logical to infer that experts who engage in different kinds of tasks in a domain will have different representations of it. However, the emphasis of research to date has been on situational or *ad hoc* goals; there has been little work done on whether long-term differences in goals lead to long-term differences in representation (Ratneshwar, Barsalou, Pechmann, & Moore, 2001). Do long-term objectives have a long-term effect on category representation? My dissertation addresses that issue, examining whether three groups of wine experts (connoisseurs, retailers, and winemakers), who have focused on objectives specific to their particular type of expertise (appreciation, selling, and production), show differences in their organization of the domain.

In a few domains, researchers have found evidence suggesting that the goals of different types of experts do affect the kind of categories they create (Medin, et al., 1997; Medin et al., 2005). Specifically, Medin et al. (1997) found that for different kinds of tree experts, specialized goals led to alternative organizations of the domain. When tree experts were given a general sorting task and asked to "put together the trees that go together by nature" there were systematic differences in the taxonomies they generated. Expert taxonomists, not surprisingly, tended to sort the trees according to scientific taxonomy. In contrast, maintenance workers showed a reliance on morphological characteristics, whereas landscapers used utilitarian categories such as whether the tree was "a good shade tree." The experts organized the same information in different ways.

The sorting differences did not appear to correspond to reasoning: "Landscapers did not appear to use their goal-derived categorize in reasoning; instead their reasoning patterns were predicted best by park personnel sorting." (Medin et al, 1997, p. 813, footnote 2). However, there were differences in the strategies used by the three types of experts.

In Proffitt et al. (2000), analysis of experts' reasoning justifications suggested some differences in their preference for different strategies with landscapers & maintenance workers using susceptibility more than taxonomists, and landscapers and taxonomists using mechanism more than maintenance workers did. However, all three groups used causal-ecological more than similarity-based reasoning strategies.

When looking at differences among types of experts, it is important to ensure that claims about differences in performance are not simply restatements of the criteria that were used to select the groups. Ideally, the different expert groups should demonstrate common knowledge bases. Medin et al. (2005) did just that in their study of fish experts. In a non-directive sort ("put the fish together that go together by nature") they found differences in the kinds of features underlying the sorts of two different expert groups. Native American Menominee fishermen relied more on ecological factors in their sorts than did the majority American culture fishermen from the neighboring town. Yet, when both groups were asked to sort the set of fish specifically according to ecological relations ("put those fish together that live together, that share a common habitat"), the majority culture sorts were not significantly different from those of the Menominee. This indicates that the difference in the initial sorts was not due to differences in knowledge—both groups knew the facts necessary to make ecological sorts. Instead, it was a matter of accessibility or preference. Whether the ecological information was consciously prioritized or simply more salient is not certain.

Practicalities

Measures of Conceptual Structure

Attempts to describe conceptual structure usually focus on the relationships among category members and the relationships among features. For between-category relationships, one key structural characteristic is overall shape. When the shape is hierarchical, the breadth (number of categories at a single level) and depth (levels of categories) are important descriptive details.

Several methods are commonly used to explore the conceptual structure of real world categories: (1) sorting tasks, (2) similarity judgments, (3) feature listing, and (4) typicality ratings. In sorting tasks, participants group items that they think are related. The resulting sorts, along with the participants' explanations of their groupings, provide information about the overall shape characteristics of a hierarchical structure and the kinds of principles that people use to organize constituents. In addition, distances between pairs of items yield an indirect measure of similarity. Direct measures of similarity (asking for judgments of how similar two items are) are useful for mapping out the relations among categories, as well as for providing an index of differentiation. Feature listing provides several kinds of information. Comparing the number and kinds of features ascribed to different categories at the same hierarchical level provides information about category differentiation, while comparing feature lists at different hierarchical levels can indicate which offers the greatest information value. Finally, typicality judgments provide insight into within-category structure and the nature of category boundaries.

The domain

The research just described has found differences among expert types, but also has found consensus. For these natural domains (fish, trees) a scientific taxonomy exists. What about a domain in which there is no single dominant default, but in which there are many cross-classifications that are equally valid, perhaps depending on the situation or the goals of the perceiver? Wine may be just such a domain.

Research on wine experts has focused on expert-novice differences with an emphasis on differences in their abilities to perceive, describe, and remember the sensory features of a wine. Little systematic attention has been paid to how wine experts organize and use their substantial knowledge base about winemaking and viticulture. My focus on how experts with different kinds of experience in the domain (as opposed to different amounts of experience) organize their conceptual knowledge distinguishes my research from the existing literature on wine expertise.

Certainly sensory input is an important aspect of the domain. Yet although taste and smell are the principal senses engaged in wine evaluation, they do not work in isolation. The first step of the formal tasting process is visual examination of the wine. Some researchers even make the provocative claim that color influences flavor perception to the extent that tasters misattribute "red" flavors to a white wine that has been dyed red (Morrot, Brochet, & Dubourdieu, 2001; Pangborn, Berg, & Hansen, 1963).

Wine expertise research has focused nearly exclusively on experts' cognitive management of sensory data, especially the perception, identification and communication of smell and taste information. Expectations about the taste profile for a particular varietal may determine which flavor characteristics they search for, subsequently identify, and ultimately remember about a wine (Hughson, 2003).

In their review of the literature, Hughson and Boakes (2001) concluded that conceptual, top-down processes based on explicit knowledge have a greater influence on expert performance than purely perceptual, bottom-up processes. At a minimum, the possession of wine tasting vocabulary—which requires focused effort to acquire—is certainly important (Hughson, 2003; Lehrer, 1975; Melcher & Schooler, 1996; Rabin & Cain, 1984; Solomon, 1990).

Other types of contextual information may also affect perception of wine (Ross, 2002). For example, researchers have found that putting a prestigious label on a bottle of "mediocre" wine results in more positive evaluations of the wine's quality (Brochet, 2001).[4] It is for this very reason that "blind-tasting" has emerged as a standard among many wine authorities today (McCoy, 2005).

Conceptual organization of wine

Gregg Solomon's (1997) research documents an intriguing situation in which experts and novices appear to have access to the same information but use it differently to organize a set of wines. In his study, a group of experts and novices tasted and described a set of ten white wines.

---

[4] Note that although this study is suggestive, it does not definitively establish that the drinkers' perception of the wine has been altered.

Solomon "scaffolded" participants' descriptions by providing them with the UC Davis Wine Wheel (Noble, Arnold, Buehsenstein, Leach, Schmidt, & Stern 1987), a tool that organizes common wine flavor terms in a hierarchical fashion. Tasters were introduced to the wheel and its structure and told to use whatever terms they thought best described the wines. Although the experts and novices did not use identical terms, both groups' selections could be well predicted by the grape varieties used to make the wines. This suggests that the experts and novices had access to substantially the same perceptual information about the wines. Despite the novices' inexperience, they were able to detect differences among the wines that reflected their composition.

A week later, participants tasted the wines again, but this time sorted them into four groups instead of describing them. The experts' groups largely reflected grape varieties, whereas the novices' sorts were based on features such as sweetness and fruitiness. Thus, when asked to make explicit groups of wines, the performance of experts and novices diverged. Experts used varietals, whereas novices did not, despite the fact that their descriptions of the same wines the prior week contained information sufficient to make a varietal-based sort.

Hughson (2003) interprets Solomon's (1997) results as evidence that wine experts use long-term memory structures based on varietal schemas. He argues that as experts learn about wines, they organize tastes and other sensations around a "mental checklist" for each grape variety. These checklists guide the taster in terms of what to look for when tasting a wine. This "Varietal Schema Theory" (VST) could explain some results reported by Hughson. Hughson modeled his study of wine experts on Chase and Simon's (1973) classic chunking experiments. Experts reviewed written wine descriptions that were either "meaningful" (i.e., the terms were consistent with a standard profile for a particular grape varietal) or scrambled (in which taste descriptors that would never be found together in nature were listed together). On a number of different memory tasks, he found that experts made more “variety errors” than novices. A variety error occurs when a term that is a legitimate description of the grape variety is offered even though it had not been presented. Another common type of expert error was the “color error.” This refers to the case in which a term commonly used to describe a wine of the same color as the target (though a different variety) is erroneously included in recall of the description. This, Hughson argued, could be accounted for by inferring that the expert had assigned the wrong variety to the sample. Experts made more of both types of errors (variety and color) than novices did, both on tasks requiring them to recall a written description of a wine and on tasks requiring them to generate their own description of tasted samples.

An alternative to VST is Global Prototype Theory (GPT), which emerged from lexical analysis of five wine tasters’ tasting notes (Brochet & Dubourdieu, 2001). From this analysis, Brochet and Dubourdieu concluded that experts’ schemas are idiosyncratic and driven by evaluative terms and “personality” descriptors as well as taste descriptors. These schemas tend to be at a more abstract level than the variety, sometimes as broad as “red wines.” Another distinction between the two approaches is that VST implies that a sample is compared to a comprehensive list of features, whereas GPT implies that it is compared to the abstracted prototype. Either theory can account for some intrusions in wine descriptions (or memories of descriptions). Yet in both, the “theories” only pertain to flavor information, failing to take into account the broader range of information that wine experts must navigate, such as knowledge of the maker of the wine, the region where it was made, and the techniques used to produce it.

Study Overview

Ross and Murphy (1999) critiqued the focus of the existing literature on concepts and categories as suffering from three limitations: “a single hierarchy, a single function, and isolated knowledge” (p. 496). This dissertation attempts to further these efforts to rectify this state of affairs.

In this exploratory study, I focus on three issues. First, how do people with extensive, rich knowledge about a complex domain organize it? Second, are there systematic and robust differences in how the different kinds of experts organize their domain knowledge? Third, assuming these differences do exist, do they affect how the experts use and extend their knowledge?

Data collection occurred over two sessions, spaced approximately six months apart. The focus of the first session was to characterize the experts and describe their conceptual organization of the domain. During this session, experts provided personal background information, completed an assessment of their expertise, and sorted wine labels into categories. The focus of the second session was to examine those categories, using multiple measures to test their psychological “reality.”

CHAPTER TWO:

CHARACTERIZING THE EXPERTS' BEHAVIOR AND KNOWLEDGE (SESSION ONE)

Introduction

One goal of the first session was to verify participants' expertise through the assessment of knowledge, background, and qualifications. In addition, I hoped to validate my initial expert group assignments. These manipulation checks are the focus of Chapter Two. The primary goal of the first session—the description of experts' conceptual organization—will not be addressed until Chapter Three. Chapter Four examines a subset of the categories generated in the sorting tasks to assess their availability and utility via primed similarity ratings, category membership ratings and an inference task.

Determination of wine expertise

Verifying that the subjects were in fact experts was essential. Though a definitive measure of expertise is elusive, researchers typically use one of four types of criteria: (1) years of experience (also known as the ten-year rule), (2) evidence of professional activity or certification, (3) word of mouth referrals, and (4) testing.

Ten years of experience in a domain has served as a common rule of thumb, or benchmark, for expert performance (Ericsson., Krampe, & Tesch-Romer, 1993; Shiffrin, 1996). But time alone is not enough; how that time is spent matters, too. On the basis of considerable research, Ericsson and colleagues argue that ten years of what they term <u>deliberate practice</u> is necessary. Ericsson emphasizes that expertise emerges not merely out of exposure to a field, domain, or activity, but as a result of intensive engagement with it.

Professional activity in a domain can be evidence of expertise. This approach works well for occupations such as doctors (Hassebrock, Johnson, Bullemer, Fox, & Moller, 1993; Norman, Brooks, & Allen, 1989), physicists (Chi et al., 1981), and psychologists (Murphy & Wright, 1984) that require extended training, but in many fields, having a job is not the same as being good at it. Furthermore, not all experts are paid for their passion, neither do all areas of expertise lend themselves to professional activity.

Another strategy is to "let those in the domain define the experts" (Shanteau, 1992), using referrals from other experts. Testing of domain knowledge or performance can also help identify experts (e.g., Gobbo & Chi, 1986).

I used a combination of these measures. Personal referrals and professional activity helped identify experts for recruitment, but I imposed additional criteria to confirm their qualifications. Analysis of domain-related activity and testing of domain knowledge provided confirmatory evidence of expert status.

Methods

Participants

Thirty experts participated in the first session, ten recruited for each of three target groups (connoisseurs, retailers, and winemakers).[5] In addition, eleven novices completed a subset of experimental tasks.[6] Seven participants were women (two winemakers, two connoisseurs, one retailer, and two novices). The average expert age was 54.3, with a range of 37 to 73 years old;

---

[5] Later testing required re-assignment of three experts, changing the proportions to 7 connoisseurs, 12 retailers, and 11 winemakers. All reported data reflect these re-assignments.

[6] During a single session lasting 15 to 30 minutes, novices completed the similarity judgment ratings (Task 3 of experts' second session), a subset of items (A1, A3, A5, B1, B2, B3, B4, C1, C2, C3, and C4) from the Background Questionnaire (Task 6 of experts' first session), and the wine knowledge test (Task 7 of experts' first session).

the novice age distribution was comparable (*M* = 52.3, range = 38 to 76 years old). The experts were fairly well educated, but connoisseurs and winemakers tended to have more formal education than retailers. Six of 12 retailers had not completed college, compared to 3 of 11 winemakers and 1 of 7 connoisseurs. Graduate degrees were the mode for the two latter groups (five each). The novices were roughly matched on the overall educational distribution: four had not completed college, and the majority some kind of a graduate degree.

Recruitment.

The procedure for recruiting experts varied slightly depending on the type of expert being approached. To recruit connoisseurs, I relied on a snowball sampling procedure using personal referrals. Starting with a few known contacts, I requested referrals for other experts in the area. Initially, winemakers and wine retailers were identified by calling names on industry lists, but I used participant referrals when possible to increase compliance. The Pennsylvania Wine Association provided a directory of wineries and I began by calling those wineries in the directory which I believed to be at least ten years old, based on the description in the directory or on the winery website. To recruit retailers, I began by speaking with staff from the Pennsylvania Liquor Control Board (PLCB) education department who provided a list of knowledgeable wine specialists from PLCB stores in my part of the state. Because I also wanted to interview retailers from outside the PLCB system, I targeted Washington, D.C., where a personal contact in wine retail led me to two participants. Cold calls to DC area liquor stores yielded two more retailers.

After identifying likely participants, I spoke with them on the phone to introduce myself and describe the study. During the conversation, I asked a few general screening questions about their background and experience. All but two participants affirmed they had at least ten years of "intensive" experience with wine. The two exceptions, a winemaker and a retailer, had eight and five years of experience, respectively.

Procedure

Location and Scheduling.

All data were collected in face-to-face, one-on-one interviews. Sessions were held in a variety of locations, including a conference room on the Franklin & Marshall campus (19% of sessions), participants' workplaces (stores or wineries, 54%), participants' homes (11%), or public spaces such as restaurants or coffee shops (17%). The first session typically lasted 95 minutes, but ranged from 45 to 180 minutes. The first session interviews occurred between April and September 2007.

General Protocol.

The initial session was my first face-to-face contact with most participants. After introductions, participants reviewed and signed a consent form authorizing me to audiotape the interview.

The experimental tasks then began. First, participants reviewed a set of 40 laminated wine labels that they would use in subsequent tasks. They then performed a preliminary sort of the labels. Once this sort was complete, they had the opportunity to perform additional sorts. The session ended with two questionnaires: one obtaining background information about the participants and a second one testing their knowledge about wine. I report on the Wine Knowledge Test, the Background Questionnaire, and the Card Review, in that order. The sorting tasks will be addressed in Chapter 3.

Wine Knowledge Test

The kinds of information used to identify potential study participants—their occupations and personal referrals—do not constitute proof of their expertise. The question remains: how to determine whether the purported experts were, in fact, expert?

Because there is no single credential that identifies wine experts, researchers have tended to rely on reputation or tests of wine knowledge as their criteria (Hughson, 2003). Hughson and Boakes (2001) developed an Australian Wine Knowledge Test that consists of eight factual multiple-choice questions about wine, plus two questions about wine-related behavior.[7] They found this test did a good job of discriminating experts (n = 28) from novices (n = 89), with mean scores of 7.25 and 2.21, respectively. The experts had perfect scores on all but four questions. In addition, no experts missed more than two questions, whereas all but one novice did.

Method

Materials and Procedure

To adapt the Australian Wine Knowledge Test (Hughson & Boakes, 2001) for American experts I removed two factual questions that were specific to Australian wines ("What type of oak is 'Grange' typically aged in?" and "What style is typical Hunter Valley Semillon?") and substituted "Chianti" for "Grange" in a third question. Because the removed items were among the four the most difficult for the Australian experts, I expected the American wine experts to score well on the test. I set a criterion of two errors as an indication that I should scrutinize potential experts more closely (though not necessarily exclude them).

[7] I incorporated the two behavioral questions into the Background Questionnaire, Task 6.

Results and Discussion

Experts scored well on the test overall. See Appendix A for the full text of the items and group averages. One third of experts had perfect scores; a single error was the mode. Experts (M = .92, SD = 0.08) performed better than novices (M = .52, SD = 0.10) on the test, $F(1,39) = 165.9$, $p = .00$, but there was no significant effect for type of expert, $F(2, 28) = .907$, $p = .42$.

Three expert participants had more than two incorrect answers. One retailer missed items 2, 3, and 6, another missed items 1d, 2, and 4, and a winemaker missed items 1d, 1h, 3, and 6. Their responses to the background questionnaire and performance on other tasks were examined carefully, but none were dropped from the study.

The wine knowledge test was successful at differentiating experts from novices. Although three experts failed to meet the experimenter-set criteria, their performance was adequate and subsequent evaluation based on other data warranted their retention in the study.

Background Questionnaire

To supplement the knowledge assessment described above, participants completed a questionnaire to assess their wine-related experience, training, and behavior. Although the Background Questionnaire provides some indication of participants' degree of expertise, the main goal was to characterize different types of experts, not qualify them for participation in the study.

An important component of the questionnaire was the identification of tasks that distinguished the expert groups. Such differences (should they emerge) could provide explanatory support for group differences in category structure and use. Of course, I also hoped the results would validate my assignment of individual experts to the winemaker, retailer, and connoisseur groups.[8]

Method

Materials and Procedure

As the sixth task of the session, participants completed the 8-page Background Questionnaire. See Appendix B for the complete instrument. The three sections of the questionnaire served different purposes. Section A measured wine consumption patterns and history. Section B assessed other wine-related training and behaviors. In addition to a number of open-ended and multiple-choice questions about wine-related education, certifications,

[8] Note that although I had tentatively identified and recruited three types of experts, it was possible that these distinctions were not the relevant ones. Instead, more—or fewer—groups might exist. Another possibility was that the groups were not mutually exclusive types. Wine expertise might be more of a progression, in which all begin as connoisseurs, but retailers and winemakers acquire knowledge about selling wine, and winemakers possess yet additional knowledge and skill sets related to production that set them further apart (Medin, personal communication). This potential overlap among the experts provided additional motivation to collect data on the types (and frequencies) of the experts' activities.

memberships, and jobs, this section included a 41-item inventory, gauging the frequency of specific activities on a 5-point scale (Never-Rarely-Occasionally-Often-Regularly). Section C collected basic demographic information. The texts of the items appear in Appendix B. Novices completed a subset of items.

Results

The main goal of the background questionnaire was to characterize experts' behavior and history with respect to wine. What—if any—kinds of differences exist among the groups in terms of their background, goals, and experience? Before exploring group differences, however, it was important to examine the behavioral data for evidence of expertise and to determine whether the initial group assignments were appropriate or required adjustment.

Expertise Confirmation

The background questionnaire contained a number of "novice indicators" intended to catch anyone who had made it through the initial screening process, but was not sufficiently expert to participate. Participants scored one point for each of the following responses. (1) If they drank wine less than once a week (c, d, or e on Item A1). No experts (but 9 of 11 novices) drank wine less than weekly. (2) If they had tasted fewer than 1251 wines in their life (e, f, g, h, i, j, or k on Item A5). Only three experts (but all novices) had tasted this few wines. (3) If they had read fewer than 3 books or articles on wine (b, c, or d on Item B3). Only one expert (but 10 of 11 novices) had read this little about wine. (4) If they claimed wine was not particularly important to them (d on Item B2). One expert (but 9 of 11 novices) agreed with this statement. (5) If they claimed neither wine-related professional activity *nor* personal involvement (a on B1 but not any of a, b, or c on B2). No experts (but all novices) claimed neither professional nor personal involvement with wine.[9]

Cluster Analyses and Group Re-Assignment

One goal of the background questionnaire was to verify the group assignments that I had made during recruitment. To do so, I conducted a hierarchical cluster analysis of the activity inventory responses (Items B8-1 through B8-41[10]), using the Ward method. This yielded two initial clusters (Figure 1). The next split, one level down, resulted in three clusters that largely match my group assignments of winemakers, retailers, and connoisseurs. See the dendrogram below for details.

Nine of the ten individuals in the bottom cluster of the dendrogram are winemakers. The initial assignment of the tenth individual (#7) to the connoisseur group was probably an error. Unlike the typical connoisseur, this individual is a wine professional—one of four connoisseur recruits who claimed some professional activity and the only one to select "winemaker" (along with two other professional roles) on item B1 of the questionnaire. In addition, his/her tasting profile (in terms of regional distribution) is much closer to that of the winemaker recruits (50% of tasted wines were Pennsylvanian, in contrast to the other connoisseur recruits who claimed only 0-1% of the wines they tasted were from Pennsylvania). Therefore, although this person's primary job was not making wine, re-assignment to the winemaker group was reasonable.

---

[9] No expert scored more than one point. One retailer, however, triggered a novice indicator *and* missed three items on the wine knowledge test. Because these were relatively small deviations and other expertise indicators (e.g., familiarity score on the card review task, described later) were consistent with the typical expert performance, this person was not disqualified.

[10] One Item (B8-38) was dropped from the inventory because it presented difficulty for the first five experts.

* * * * * * * H I E R A R C H I C A L   C L U S T E R   A N A L Y S I S * * * * * * *

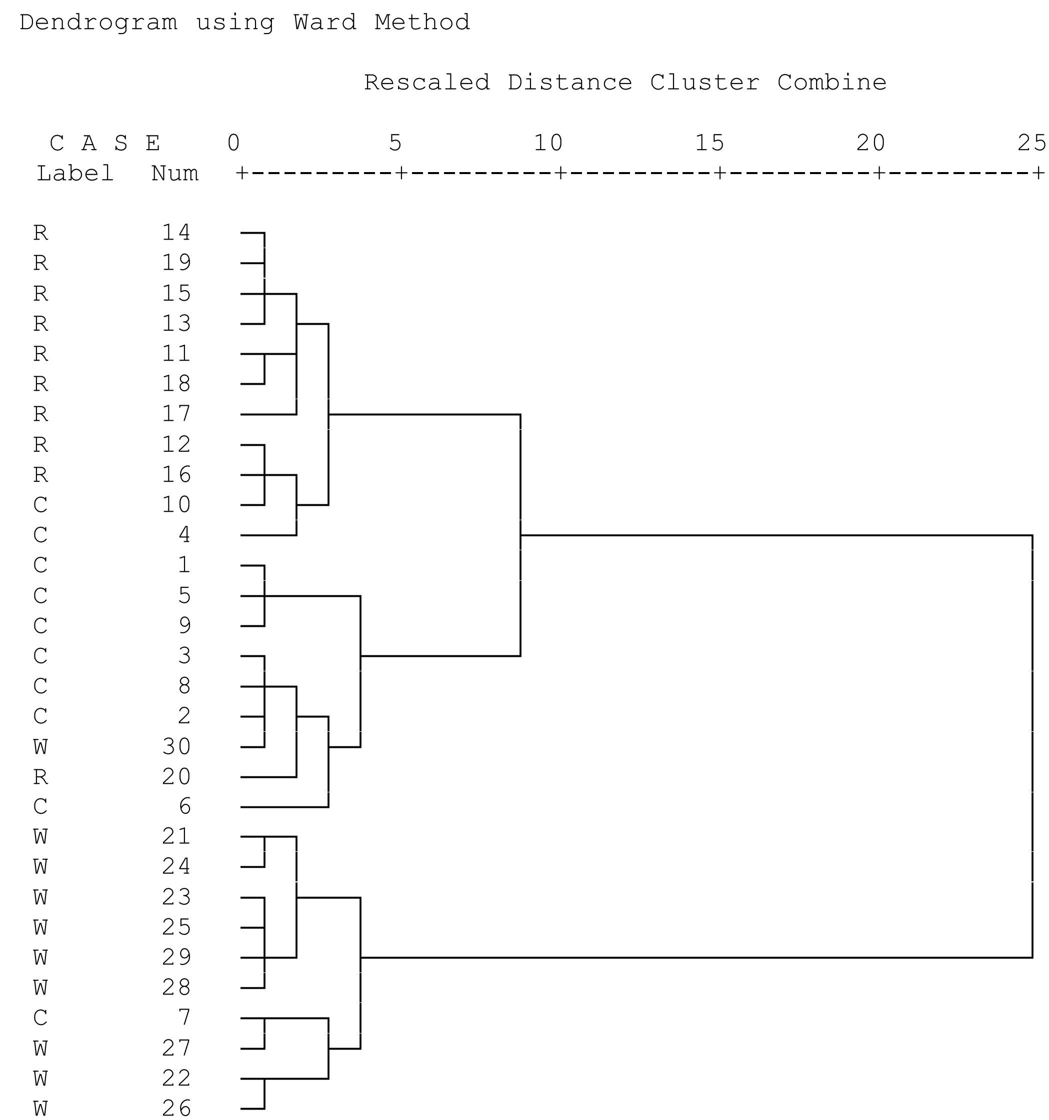


Figure 1. Expert clusters based on activity responses

Note. The labels (C/R/W) indicate the tentative group assignments at recruitment.

At the top of the dendrogram is a cluster containing nine retailer recruits and two connoisseur recruits (#10 and #4). The initial classification of these two connoisseur recruits had been uncertain. Although they did not work in a wine store, they had some professional involvement with the domain. Neither had claimed “no professional involvement” on item B1. And while neither had selected the “wine retail” option, both were involved with restaurants in some capacity. I reassigned both to the retailer group.

The last of the three clusters consisted largely of connoisseur recruits (7 of 9), plus one retailer and one winemaker. The winemaker (#30) was retired, which may account for the low scores on the winemaker-type activities. S/he had, however, owned and operated a winery for at least ten years. Though these activities were in the past, they were the source of his/her expertise, so I left this expert in the winemaker category. The retailer recruit (#20) was not a prototypical retailer (on Item B1, s/he had selected neither “commercial wine sales or marketing” nor “wine retail,” but had chosen “sommelier or wine steward” along with several other professional activities). Because this person is a professional whose career focuses on the domain, I did not re-assign him/her, but kept the original retailer category.

These re-assignments resulted in unequal groups: 7 connoisseurs, 12 retailers, and 11 winemakers.

Descriptive Statistics

Using the new group assignments, what were the characteristics of the experts overall, and did they differ among expert groups?

Drinking Behavior (Section A)

Most experts (70%) reported drinking wine daily (Item A1). A two-way chi-square analysis did not find a relationship between expert type and choice of frequency category, χ^2.(2, N = 30) = 1.346, p = .510. Estimates of personal wine cellar size (Item A2) showed considerable variability among experts (ranging from one case to more than 4500 bottles), but expert type did not predict the size of the cellar, F (2,26) = 1.18, p = .321. There was also variability in the number of different wines individuals tasted annually (Item A3), with retailers having the greatest exposure to different wines.[11] All experts had at least five years of experience (Item A4) tasting wines at these rates (range = 5 to 40 years), with an overall average of 17.8 (SD = 9.9); there were not significant differences among the expert groups. Tables of drinking patterns are in Appendix C. All but three experts estimated that they had tasted at least 1250 different bottles of wine over the course of their life (Item A5); there were not group differences.[12]

Description of the geographic distribution of recently tasted wines suggests some differences in the range of wines tasted. Pie charts of the regional distributions appear in Appendix C. There were significant effects of expert type on the percent of Pennsylvania, New York, and "other" wines consumed, Fs(2,27)=20.16, 6.164, and 6.787, respectively, ps <05. Post-hoc tests established that winemakers consumed more of these less mainstream, predominantly east coast, wines than either of the other two groups (Tukey's HSD, ps < .05). There were not significant group differences in the consumption of wines for the other ten regional categories.[13]

---

[11] For this analysis, I collapsed responses into two categories: more or fewer than 500 wines per year. No connoisseurs claimed they had tasted more than 500 wines, but 67% of retailers and 36% of winemakers did, χ^2.(2, N = 30) = 8.283, p = .005.
[12] For this analysis, I collapsed responses into two categories: more or fewer than 1250 lifetime wines. All winemakers, but only 71% of connoisseurs and 92% of retailers, had tasted more than 1250 wines in their lives. Still, there was not a significant relationship between expert type and this dichotomous grouping, χ^2.(2, N = 30) = 3.942, p = .120.
[13] The results suggested marginal effects for California, F(2,27) = 3.073, p = .063, France, F(2,27) = 2.418, p = .108, Italy, F(2,27) = 2.81, p = .078, Spain, F(2,27) = 2.309, p = .119, and Australia, F(2,27) = 3.258, p = .054.

Education and Experience (Section B)

Winemakers had the most formal wine-related education (Item B6), while nearly all retailers and connoisseurs were self-taught. Degree of personal involvement with wine (Item B2) discriminated novices from connoisseurs, but was not that relevant to the professionals for whom wine was not a hobby, but either a life-altering passion or "just a job." Connoisseurs were the only expert participants to claim "no professional involvement" (Item B1). Most of the retailers selected wine retail, as expected, with nearly half of them checking "commercial wine sales or marketing" or "sommelier or wine steward." All winemakers described themselves as winemakers. In addition, half were involved in wine retail and nearly half claimed some type of "other professional activity."

Activity ratings (Section C)

In the inventory of wine-related activity, experts indicated how frequently they engaged in a variety of activities. See Appendix D for group means on each item. These ratings served as the input for the cluster analysis described above. In this section I examine which activities distinguish the groups from each other. To assess this, I conducted a principal components factor analysis on the 40 inventory items. See Appendix E for the scree plot and a table of the factor loadings for components 1 through 5. I will focus on the first three components, which accounted for 61.86% of the variance.

The loadings on the first factor were positive and fairly large, with one exception. Item B8-13 ("Study wine rankings and/or ratings") received the lone negative loading (-.126). Examination of the mean scores reveals that this was an unusual item, in that it was one of only two statements for which connoisseurs had the highest score,[14] and the only one on which the retailers' mean score was meaningfully higher than winemakers' (*M*s =4.00 and 2.91, *SD*s = 1.414 and 1.221, respectively). Thus, this first component seems to identify the one behavior that connoisseurs are more likely to pursue: the study of wine rankings and ratings.

On the second factor, there were fifteen negative loadings. These scores clearly identified winemaker-specific activities. The following items all received 2nd factor loadings of -.2 or less: (B8-9) plant or tend grapevines, (B8-10) harvest grapes, (B8-11) taste grapes, (B8-18) decide when to harvest grapes, (B8-19) decide what barrel type to use, (B8-23) sell wine in bulk, (B8-30) think about disease prevention for grapes, (B8-33) think about pest management, (B8-34) sample wines before they are mature, (B8-35) use heavy machinery, (B8-36) research new wine making procedures, (B8-40) conduct a sensory evaluation of a wine, and (B8-41) conduct a chemical evaluation of a wine. For those items with factor loadings near zero, winemakers and retailers (but not connoisseurs) had roughly similar mean scores. These items were (B8-14) serve wine to customers, (B8-15) set wine prices, (B8-20) sell wine by glass, (B8-24) taste wine for flaws, and (B8-37) manage inventory.

The third factor loadings identified items that characterized retail behavior. The two items with the highest loadings were (B8-21) sell wine by the bottle and (B8-37) manage inventory. Other items with positive loadings greater than 0.2 were (B8-14) serve wine to customers, (B8-15) set wine prices, (B8-20) sell wine by the glass, (B8-27) help others choose wines to buy, (B8-28**)** encourage others to buy a particular wine, **(**B8-39) and plan special wine

[14] The other, Item B8-16, "Think about wine-food pairings for self and family" had the next lowest factor loading (.237). Although connoisseurs had the highest score (M = 4.71), it was not very different from the scores for the other groups (*Ms* = 4.42 and 4.64 for retailers and connoisseurs), respectively.

promotion events. Retailers had relatively high means on all of these items, but were not alone in that. Often, winemakers (and occasionally connoisseurs) also engaged in these persuasion-oriented activities.

Discussion

The background questionnaire provided confirmatory evidence for the expertise of the expert participants. Their performance on designated "novice indicators" was clearly distinct from that of the novices. Cluster analyses of the activity inventory largely confirmed the recruitment groups, but led to the reasonable re-assignment of three connoisseurs to other groups.

Reports of drinking behavior identified two differences among the three groups. First, retailers tended to taste a larger number of different wines on an annual basis. Second, winemakers drank higher percentages of local (east coast) wines than the other two groups.

Factor analysis of the activity inventory found only one activity that connoisseurs engaged in more than the other expert types—the study of wine ratings and rankings. A group of activities clearly related to the process of making wines set winemakers apart, but the second factor loading also identified a cluster of consumer-oriented activities that contribute to the job description of both winemakers and retailers. The third factor identified sales and promotion activities that largely characterize retailer behavior.

Card Review

The primary objective of the card review was to prepare participants for the subsequent sorting tasks by familiarizing them with the stimulus set. This pragmatically necessary task also provided supplementary data on level of expertise.

Method

The set of 40 wines that participants reviewed serves as the basis for most of the tasks in this dissertation. A detailed description of the selection process follows the description of the materials and procedure.

Materials

The front and back labels (as well as any collars) of 40 wine bottles were glued and laminated onto an 8 x 5 inch piece of white card stock. Most of the labels came from an on-line database maintained by the Alcohol and Tobacco Tax and Trade Bureau (TTB). This database contains pdfs of the Certificate of Label Approval (COLA) applications required for most containers of alcohol sold in the United States. Because the applications include copies of the labels, this was a very useful resource. For most of the 40 wines, I downloaded the applications from the database, and then reduced, color printed, and trimmed them. For those wine labels not available from the database, I obtained labels from the wineries (3 wines) or purchased a bottle and removed the label (5 wines). As necessary, these were reduced on a color-copier to fit on the card. In addition to the label, each card had a randomly-assigned tracking number (between 1 and 40) in the upper right hand corner. See Appendix F for thumbnail images of the labels.

Procedure

Participants received the stack of 40 wine label cards (shuffled and presented in a random order) and heard the following instructions:

> I am interested in how experts organize their knowledge about wine. Each of these cards has the label from a bottle of wine on it. I'd like you to go through the set of cards, and for each one answer some questions.
>
> First, are you familiar with that wine?
>
> Second, have you ever tasted a bottle of that particular wine?
>
> Finally, I would like you to tell me two important things about the wine. Note that even though you may never have tasted this particular wine, you may still be able to say some things about it with high confidence.

The full instructions were repeated for the first few items and then only as needed. I used audiotape and notes to record responses.

Stimulus Selection

The selection of the specific set of wines to use presented a significant challenge. Hundreds of thousands of wines are produced each year and available for sale in the United States. To make the task manageable, I limited the set to the more reasonable size of 40 wines. The question was, how to select 40 out of such a large set?

There are a number of dimensions that could be used to organize the domain of wine. In selecting the wine sample for this research, I paid attention to region, color, price, source (vineyard, winery, importer, or bottler), vintage, and varietal (grape). I sought wines that varied (but could also cluster) on these key dimensions, making an effort to minimize the correlation among them. To do so, I designed my sampling procedure to ensure representation across combinations of general categories for the first three dimensions, crossing five broadly-defined regions with three wine color categories, and also with three price groups. I also checked that there were at least some clusters of the last three variables (small sets of wines from the same source, of the same vintage, and from the same varietal). Again, the goal was to ensure satisfactory representation of important wine categories and break correlations between key variables.

Sampling Procedure.

The Pennsylvania Liquor Control Board (PLCB) is the largest purchaser of wine and spirits in the United States. I obtained two PLCB product lists, the "Regulars" (standard products available in basic stores across the state) and the "Specials" (specialty items that are available in limited supply, and only at selected premium stores). After editing these lists to remove non-wine items (e.g., hard liquor), wines sold in sizes above 1.5 liters, and non-traditional or "borderline" wine products (e.g. wine coolers), 13,562 wines remained. This set—all the wines available for sale in Pennsylvania through the PLCB stores during early 2007—constituted my initial working population.

I used a stratified sampling procedure combined with a partial quota system. A simple random sample would have been inappropriate, because of an asymmetry between the two product lists. There are 5214 Regular items, compared to 12,066 Specialty items.[15] However, the larger Specialty list accounted for less than 4% of PLCB business in 2005-2006. To correct for this disparity, I stratified the sample and adjusted the weights for the two lists to create a sub-sample of 200 wines, consisting of 60% Regular items and 40% Specialty items. Given the sales

---

[15] These numbers reflect the full range of products, including hard liquor, wine coolers, and other non-wine products.

figures, it may seem that I overemphasized the Specialty list. However, because the sales figures include a variety of products, they need not be taken as a strict guideline with respect to wine. Moreover, the Regulars list contained a high proportion of non-vintage, mass-market wines; without the shift in weights these would have dominated the final wine set.

Quotas.

To ensure minimal representation of a good cross-section of the regional, wine type, and price categories, I established two quota grids. (See Appendix G.) The first crossed five regional categories with wine type. Thus for each of five regions I obtained a minimum of four wines: one each of the three types (Red, White, Other[16]) plus a second wine of any type. The second quota grid crossed the same regional categories with price. I sought one entry for each of the 35 cells. Wines pulled at random from the 200-wine sub-sample could fill a cell in one or both quota grids, depending on what was open.

The goal of the quotas was to ensure that there was some representation from each of the five regions of interest and that no single type of wine or price level would be exclusively associated with a given region (and vice versa). The quota system ensured minimal representation, but did not dictate the overall distribution of wines for the set.

Wines that could not fill a quota cell were put on a waitlist. If no wines from the sub-sample could fill a cell, I returned to the full working population to find suitable wines. This was necessary for two cells: New World (Non-US) X Other Type and New World (Eastern) X High Price.

---

[16] "Other" included pink, blush, and rose wines as well as other specialty wines such as dessert or sparkling wines.

The process yielded 25 wines, which I next reviewed for clusters of two or more wines sharing key properties. Specifically, I wanted to ensure that the final set of wines contained at least three clusters of wines (a) of the same vintage, (b) made from the same grape variety, and (c) from the same source (vineyard, winery, importer, or bottler). Sufficient clusters existed for vintage and grape variety. To create source clusters, I added three wines from the 200-wine sub-sample that obviously shared a source with wines in the existing set of 25.

The first 12 wines from the waitlist completed the set of 40 target wines. Despite significant effort to obtain labels for these exact wines, it was not always possible, so I sometimes substituted close, but slightly different labels (e.g., a different vintage or product from the same producer). See Appendix F for images of the final set of 40 wines; Appendix H presents the names of the wines in a table of familiarity scores and Appendix J characterizes each of the wines on some key dimensions.

Results

The primary goal of the card review was to familiarize participants with the wines in the set. In addition, this provided information about how well participants knew these specific wines and whether there were group differences in their experience with the wines. The task also provided an index of the wines, identifying those that were more and less familiar, helping target wines for inclusion in future tasks. See Appendix I for comments on the procedure and results.

Individual & Group Scores

Familiarity.

Table 1 presents the familiarity and tasting results. On average, experts were familiar with 58% of the wines. A one-way ANOVA examining the impact of expert group on familiarity found a significant effect of expert group, $F(2,1197) = 18.24$, $p = .00$. Retailers were more familiar than both connoisseurs and winemakers (Tukey's HSD, *p*s = .00). In addition to having the four highest individual scores (78% to 95%), all but one retailer claimed familiarity with more than half of the wines. The two lowest-scoring retailers (43% and 53%) were from the D.C. area.[17] The average score for connoisseurs was 51%, ranging from 30% to 70%. The average score for winemakers was also 51% (ranging from 15% to 70%), but the two lowest scorers for the entire set of experts were both winemakers (14% and 25%).

Table 1. Average wine familiarity and tasting experience by expert group

Mean (St. Dev.)	Connoisseurs $n = 7$	Retailers $n = 12$	Winemakers $n = 11$	Overall $N = 30$
Familiarity	.51 (.501)	.68 (.467)	.51 (.501)	.58 (.494)
Tasting	.30 (.459)	.46 (.499)	.31 (.465)	.37 (.465)

Tasting.

Although tasting results were lower than the familiarity ratings, they followed the same general pattern. Expert group predicted tasting experience in a one-way ANOVA, $F(2,1197)$, $p = .00$. Post-hoc tests established that retailers had tasted more of the wines than the connoisseurs and winemakers (Tukey's HSD, *p*s = .00). Retailers had the highest average score (46%) as well as the eight highest individual scores (ranging from 45% to 68%). Connoisseurs had the lowest

[17] The relatively high proportion of east coast wines (as dictated by the quota system) may have reduced the performance of D.C. retailers. In addition, the fact that the wines were drawn from those available for sale in Pennsylvania certainly should have increased the chance that PLCB employees (three of the ten retailers) knew them. Note, however that store contents vary and some wines were available only online or by special order.

average score (*M* = 30%, range from 15% to 43%), but winemakers were close (31%, range from 13% to 43%) and had the two lowest individual scores (both 13%). Because the tasting results do not add much information to the familiarity results, I do not include them in subsequent analyses.

Item Analysis

Examining the average familiarity scores for individual wines helped identify wines that should be excluded from tasks in the second session. Wines that at least 75% of each expert group had rated familiar were good candidates for inclusion in future tasks; those that 75% had rated as unfamiliar would best be omitted. By this calculation, the number of familiar and unfamiliar wines varied by group, echoing the overall pattern reported above. Retailers were familiar with the most wines (18), followed by connoisseurs (13) and winemakers (10) and unfamiliar with the fewest (2, compared to 13 for connoisseurs and 11 for winemakers). A table reporting familiarity by individual item appears in Appendix H. The familiar wines were all French or American, whereas the less familiar wines were from New World countries outside the U.S. (South Africa, Argentina, Australia) or non-French Old World countries (Italy and Portugal).

To what extent were the groups consistent in terms of the wines they knew? It could be that the different expert groups had familiarity with different wines. For example, the “classically-trained” connoisseurs may have disregarded wines from the eastern U.S., whereas one might expect winemakers to know local wines well. Some differences emerged in the wines known best by each group. Despite their generally lower familiarity results, there were two wines that winemakers knew better than the other two groups did; these were both from Pennsylvania.

Discussion

Overall, performance on the card review task was lower than expected, but not easily avoided given the immense range of wines available in the marketplace. Identifying those wines that were most familiar to participants was an important preparatory step for the second session. Retailers stood out as having greater familiarity and experience (tasting) the wines in the set. It is possible that their superior performance is due to the source of the wine sample—the PLCB product list that many (though not all) of the retailers work with daily. On the other hand, their familiarity with a broad set of wines may be due to the nature of the work—remember that on the Background Questionnaire, they reported broader tasting experience than the other groups.

Summary

This chapter reviewed three tasks from the first session —the Wine Knowledge Test, the Background Questionnaire, and the card review. These three tasks were successful in establishing the expertise of participants, ensuring that their group assignments were appropriate, identifying the most (and least) well-known wines in the set, and describing some key characteristics of these three expert groups.

Retailers evidenced broader exposure to and familiarity with wines, as well as regular engagement with customers and the management of wine inventory. Connoisseurs demonstrated greater focus in their wine drinking behavior, an interest in ratings and rankings, and few activities that distinguished them from the other expert type. The winemakers had the most exposure to east coast wines, the highest degree of formal wine-related training, and two clusters of typical activities. The first cluster, that set them apart from other experts was production oriented. The second set overlapped with retailers and described the customer-oriented component to their respective livelihoods.

In the next chapter, we examine whether differences emerge in the categories the three expert groups use to organize the domain.

CHAPTER THREE: CATEGORY ELICITATION (SESSION ONE)

This chapter presents the sorting tasks, which were used to identify the types of categories and conceptual structures the three expert groups used to organize the domain.

Having familiarized themselves with the stimulus set, participants next performed a progressive sorting of the wine label cards. In addition to creating hierarchically-related groups of wines, they also gave each group a name, explaining why the wines belonged together. Thus, this task provided information not only about the structure of their conceptual organization of the domain, but also about the principles underlying that structure.

Because one of the objectives of this research was to test the hypothesis that experts had multiple conceptual organizations available, it was important that experts have the opportunity to revisit the set of wines and sort them more than once. Informing participants that they would be able to complete multiple sorts was intended to reduce pressure to incorporate multiple dimensions into a single sort.

It may seem odd to include a hierarchical sorting task when one motivation for the research was concern that a single hierarchical taxonomy is insufficient to describe a domain's conceptual organization. However, although the sorting technique used here licenses hierarchical sorts, it does not mandate them. The participants always have the option of declining to proceed with splitting and clumping. Indeed, when fish experts completed a similar task, some Menominee fishermen created only a single-level sort (Medin et al., 2003). Some tree experts (ecologists) have had difficulty creating mutually-exclusive groups at a single level (they wanted to assign trees to multiple categories) (Medin et al., 1997) but that type of problem did not arise for the wine experts.

Designing a study that did not allow participants to create hierarchies would have stacked the deck against them. My hypothesis was not that no hierarchical organization exists, but I did not expect it to be the only organization.

My observations and intuitions about the domain led me to expect that the experts would have multiple ways of organizing it available to them. Given Murphy and Ross's (1999) findings with food categories—that items could be cross-classified into taxonomic or script categories—I expected that given the opportunity experts would cross-classify wines as well. These might well take the form of taxonomic and script categories, as Murphy and Ross had observed. Certainly wines differ in their appropriateness for different situations.

The literature on wine expertise leads to some predictions about the kinds of characteristics and categories that will be important to experts. Varietal-schema theory predicts the use of varietal-based categories (Hughson, 2003; Solomon, 1997) by experts. Global prototype theory suggests that color, color-based flavor terms and evaluative information may be common among experts (Brochet & Dubourdieu, 2001).

The broader expertise literature is less clear on what to expect as differences among functional types. The limited research among different expert types suggests that utilitarian or script categories may emerge. Retailers might identify wines preferred by a particular type of customer or for particular events. All experts may be sensitive to the context in which a wine will be drunk, e.g., the food pairing or role during a meal as a dessert wine or aperitif. Winemakers may attend to differences in the production of a wine. Connoisseurs may be concerned with the value or status of a wine.

The background questionnaire revealed group differences in the range of wines sampled by the different groups as well as some differences in their wine-related activities and behaviors. Based on these differences, one might expect the three groups to emphasize different types of categories and features in their sorts. In terms of reported drinking behavior, retailers demonstrated the broadest exposure to different wines. Retailers had geographically diverse tasting experience in contrast to winemakers' relatively narrow, local focus. These differences in exposure may result in a stronger emphasis on regional categories for retailers relative to winemakers. The activity inventory identified a number of behavioral differences. As consumers who "study wine rankings and ratings" the connoisseurs may emphasize quality evaluations more than the two other groups. Given winemakers' occupational focus on wine production, they are the experts most likely to use process-oriented categories. In particular, they may be more sensitive to varietal-specific idiosyncrasies and style profiles. However, because the details of the winemaking process can have a direct impact on a wine's taste and are common fodder for reviews and other wine writing, I would not be surprised to see some use of process-oriented categories among connoisseurs and retailers as well.

All three groups were comparable in the frequency with which they conducted sensory evaluations of wines, attended and conducted wine tastings, gave verbal descriptions of wines, and thought about wine-food pairings for themselves, so I would not predict big differences in their emphasis on taste or style characteristics of wines.

Sorting Task

The three main goals of the sorting tasks were (1) to determine whether experts had multiple domain organizations available, (2) to describe them in terms of structure, content, and item distances, and (3) to explore group differences in the sorts.

Method

Materials and Procedure

Preliminary sort.

The basic procedure followed Medin et al.'s (1997) approach with tree experts. Participants received the full set of 40 wine label cards[18] and heard the following instructions:

> Now look over the wines you have just reviewed and put together the wines you think belong together, creating as many or as few groups as you need. There may be many ways to think about organizing this set of wines, and I will provide you with an opportunity to explore other possibilities. For now, however, just use whatever groups seem most natural to you.

After participants had created their initial groups of wines, I asked them to provide a name or label for each group.[19] After recording their answers, I encouraged them to "split as many of these groups as you'd like into smaller groups."

I recorded the subgroups and requested labels as before, repeating the process of (1) inviting them to split groups, (2) requesting labels, and (3) recording results until they were done subdividing the groups. I then restored the groups of wines they had created at the initial level and asked them to "put together any groups of wines that you think belong together." I recorded which groups were consolidated and requested labels for these new, larger groups before

[18] This represents a slight deviation; Medin et al. (1997) omitted items unfamiliar to individual participants.

[19] This first set of categories will be called the "initial level" to distinguish it from the first complete, multi-level sort, which will go by the term "first sort."

repeating the process of (1) inviting them to clump groups, (2) requesting labels, and (3) recording results, until the participant was done.

Subsequent sorts.

Except for the introduction and conclusion, the procedure for subsequent sorts was the same as for the first sort. As introduction, I said:

> You just finished organizing this set of wines into a hierarchy of groups. Now I'd like you to think about whether there is any other way of organizing them that is also meaningful to you. Can you think of one?

If so, I continued as follows:

> Great. We'll follow the same procedure as before, starting with an initial set of groups and then splitting and lumping them as you see fit. Please start by creating a set of initial groups, putting together the wines you think belong together. Use as many or as few groups as you need.

When they finished splitting and clumping the wines for their second sort, I prompted again for another complete, novel sort, by asking, "Can you think of any other meaningful way of organizing these wines?" and repeated the process. At the end of the third sort (or earlier, if time appeared to be limiting their willingness to proceed), I used the following script:

> I won't ask you to complete any other full sorts of these cards, but I'd like to know if you can think of any other meaningful ways of organizing the domain. If so, can you briefly describe them?

All of the experts, but none of the novices, completed this task.

Results

How did experts organize the domain and did they have more than one conceptual organization available? To answer these questions, I examine the structural characteristics of the sorts (e.g., their breadth and width), the similarity relations the sorts reveal, and the explicit content of the sorts based on the labels the experts provided. Running in parallel to these analyses is the question of whether there were systematic differences in the kinds of sorts the three expert groups created.

General Sort Descriptives

Did experts have access to more than one conceptual organization of the domain? The 30 experts completed 70 full sorts, averaging 2.33 sorts per person (*SD* = .758). Half of the participants generated the maximum number of sorts allowed by the task, and of these two-thirds (10) described additional meaningful ways they organize the wines. Two connoisseurs and three retailers limited themselves to a single sort; of these, one connoisseur and two retailers offered no additional descriptions. This evidence supports the intuition that experts have more than one way of organizing the domain.

The basic structural characteristics of a sort are its depth and width. Differences among the expert groups in terms of the average number of levels per sort and average number of categories per level were neither expected nor observed.[20]

Similarity Structure

Which wines did experts view as more and less similar? Did the groups differ in their perception of the relationships among the wines? The sorting data permits several ways of answering these questions by providing a distance measure for every possible pair of wines.

---

[20] I operationalized depth as the total number of levels in a sort and width as the number of categories per level. Over the full 70 sorts, the number of levels ranged from 1 to 8, with an average depth of 3 levels (mode and median; *M* = 3.01, *SD* =1.49). In terms of width, the number of categories per level ranged from 1 to 25, with a mean of 7.73 (*SD* = 5.892, median = 6, mode = 2). There was large variation in the number of categories generated, with individuals creating between 14 and 219 per person (*M* = 54.4, *SD* = 39.8, median = 46) and between 2 and 74 per sort (*M* = 23.3, *SD* = 17.8, median = 19).

Each sort yielded a pair-wise distance matrix indicating the distance between each pair of wines. The distance value is equal to the number of levels that must be crossed to put two wines together in the same group. Thus, all wines have a distance of 0 from themselves. Two wines that are together at the most specific level will have a distance of 1, and so on.[21] A low distance score represents two closely related wines. Over the 70 sorts, this resulted in more than 54,000 distance scores. My objective is to determine (1) what (if any) clusters of wines emerge from the experts' sorts and (2) what (if any) clusters of experts exist based on these sorting distances.

Multidimensional scaling of wines.

To examine the perceived relationships among the wines, I used the non-metric MDS ALSCAL procedure (SPSS 16.0.2 for Macintosh, 2008) selecting the two-dimensional default. This created an un-rotated plot of the Euclidean distances among the wines (Figure 2) based on the pair-wise wine distances derived from experts' first sorts. The input was a 40 x 40 distance matrix of the wines, created by averaging the distance scores for each pair of wines across all experts' sorts. The fit of this multidimensional scaling solution fit to the average distance matrix (RSQ = .72) is considered acceptable (Garson, 2008).

---

[21] To generate these distances from the sorting responses, I used free software developed by NSF Award 0527707 to Norbert Ross, Tom Palmeri and David Noelle.

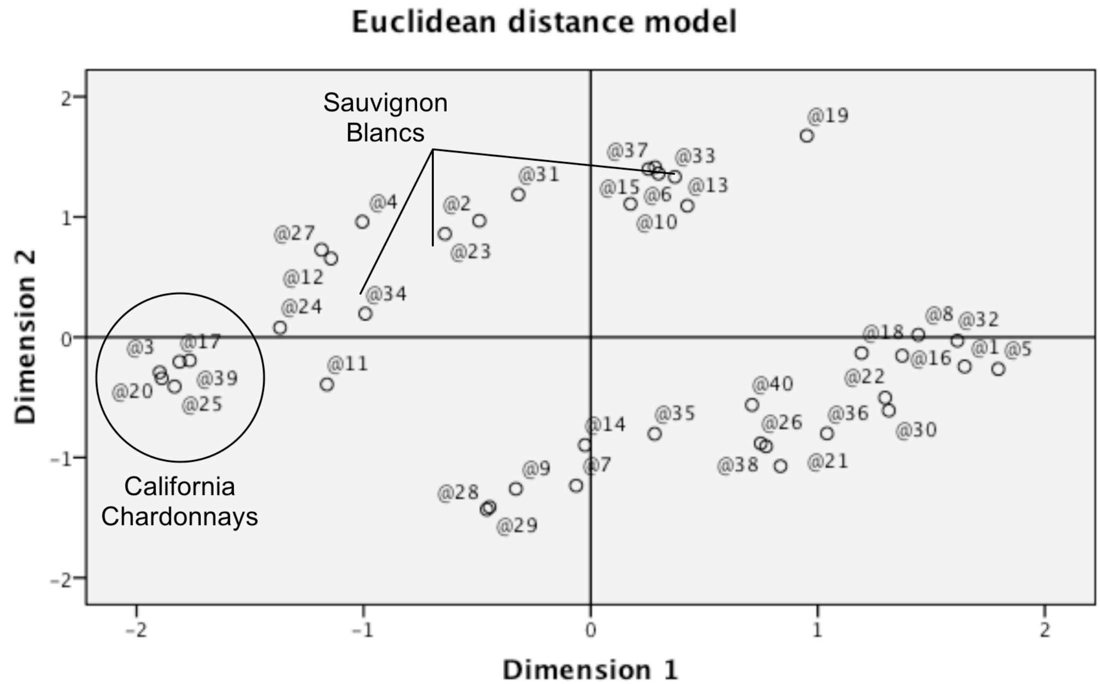


Figure 2. Multidimensional scaling of the distances among the wines: All experts

Note: The numbers identify the wines. See Appendix J for a list of the wines along with key characteristics. To obtain this plot, the SPSS ALSCAL procedure mapped the wines to approximate the distances provided by the averaged distance matrix of all sorts for all experts.

The plot demonstrates the importance of color and region in the experts' sorts. On the chart a diagonal strip clearly separates the reds (bottom right) from the other wines, making color a reasonable interpretation of Dimension 2. Dimension 1 reflects a geographic distribution, though not a perfect west-to-east map. Progressing from left to right within the two color swaths, the regions progress roughly as follows: California, other U. S., other New World (Australia, South America & Africa), Old World (Europe). Grape varietal clusters occasionally appear. For example, the cluster of five wines at the left side of the chart (3, 17, 20, 25, and 39) are all California Chardonnays. It seems, though, that this varietal cluster is a function of the coincidence that nearly all of the California whites were chardonnays. When region and grape are pitted against each other, grape loses. For example, the three sauvignon blanc wines (2, 33, and 34, from New Zealand, France, and Washington state, respectively) are more dispersed. See Appendix J for a list of the numbered wines sorted by color, region, and grape varietal.

When the distances were averaged for each expert group separately, the resulting plots for retailers and winemakers were roughly comparable to the overall plot shown in Figure 2 above (see Appendix K). For connoisseurs, however, the clear separation of reds and whites and the regional progression were muddied (Figure 3).

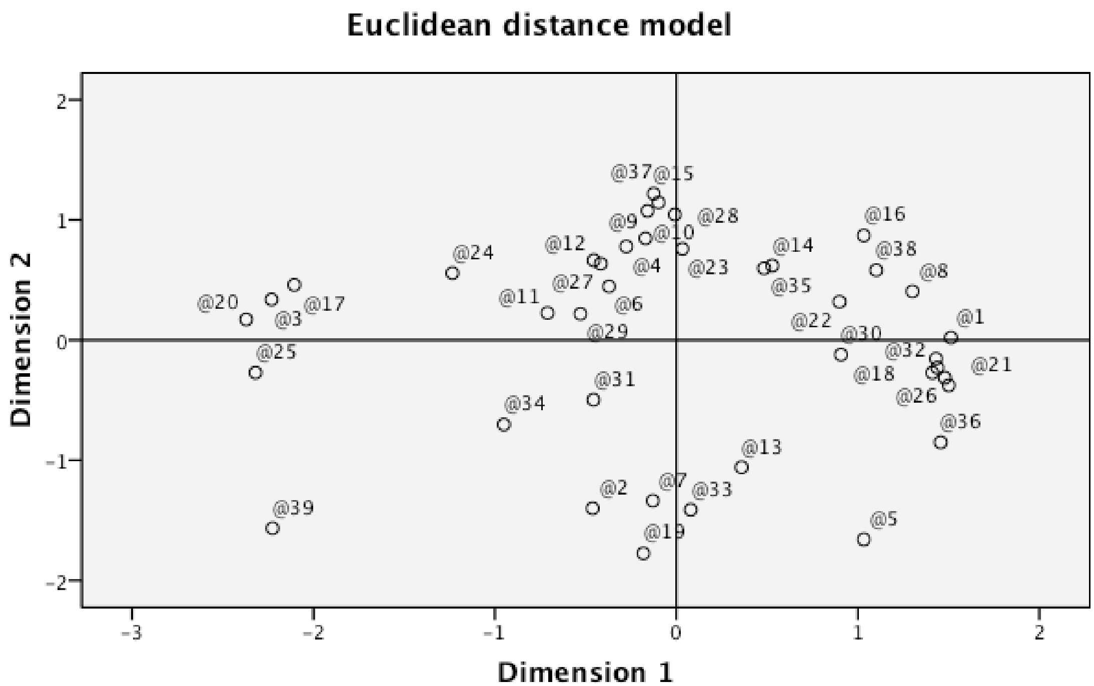


Figure 3. Multidimensional scaling of wines: Connoisseurs

Note: The numbers identify the wines. See Appendix J for a list of the wines along with key characteristics. To obtain this plot, the SPSS ALSCAL procedure mapped wines to approximate the distances provided by the averaged distance matrix of all connoisseurs' sorts.

For connoisseurs, there are more whites on the left half of the plot and more reds on the right, but there are some exceptions that are not easily explained. Three California reds (9, 28, and 29) appear in the large cluster of white wines in the upper left quadrant. Two red wines (5 and 7) are scattered among the whites towards the bottom center of the plot. It is tempting to claim that this mixing has occurred because of an emphasis on price or quality. The four most expensive wines in the set (19, 7, 39, 5) are arrayed along the lower edge of the plot. However, the next four wines in price (40, 4, 35 and 32) are not particularly close by. The wines most despised by the connoisseurs are not at the top of the plot, but are relatively close to the x-axis. (To protect the wines' feelings, I do not identify them by number.) Neither is the problem that the scaling was a poor fit to the connoisseurs' average distance matrix (RSQ = .703).

Distances for all expert types.

To assess the overall level of consensus among participants and to identify expert subgroups I used Romney et al.'s (1986) Cultural Consensus Model (CCM), a factor analytic technique.[22] A good fit of the model indicates strong agreement among participants and is considered evidence of underlying consensus. The three criteria the model stipulates for a good fit are (1) positive loadings on the first component, (2) a relatively large 1st eigenvalue that accounts for several times the variance of the 2nd eigenvalue, and (3) that subsequent eigenvalues are all small and "diminish slowly in size" (Romney et al., 1986, p. 323). Researchers have interpreted a good fit of the model with certain patterns of residual agreement as indicative of

---

[22] Weller (2007) distinguishes between a formal and informal version of the CCM. The formal model has more stringent assumptions about the data collection process and permits estimation of correct responses. The informal model is more lenient in its restrictions and does not result in a true "competence score" for each subject in terms of a measure of their knowledge. Instead, the competence score reveals their correspondence with other respondents. I use the informal model here.

generally shared expertise, with differences due to subgroups (e.g., Bailenson, Shum, Atran, Medin, & Coley 2002; Boster & Johnson, 1989; Medin et al., 1997; Medin et al., 2005; Shafto & Coley, 2003).

For each expert's first sort, I conducted a principal components factor analysis on the 780 distance values (one for each possible pair of wines) using expert as the variable. This analysis examines the degree to which underlying factors account for correlations in an expert-by-expert similarity matrix derived from comparison of experts' pair-wise wine distances. The model did not fully meet Romney et al.'s (1986) criteria; the fit was marginal. Although all experts had positive loadings on the first factor, the first eigenvalue (9.36) accounted for 31.19% of the variance, only 2.7 times that of the second value, just shy of the recommended 3 to 1 ratio (Weller, 2007). Subsequent eigenvalues were relatively small and within a narrow range.[23] When the CCM does not reveal a strong single factor solution, as in this case, Weller (2007) advises

---

[23] Concern that the distances were misleading spurred a re-analysis using binary measures. This concern stemmed from the following scenario: Imagine an expert whose first sort uses regional categories to group wines at the initial level and then subdivides the regions by varietal. For the second sort, s/he begins by creating varietal-based categories at the initial level and then subdivides those by region. The two sorts take different routes, but arrive at the same final categories. Although the sorts are certainly not identical, but the principles have not changed, just their ordering. However, distance matrices from these two approaches would look quite different (and the differences would compound as the number of levels increased). This phenomenon made me question the value of the pair-wise distance values as input for the determination of the amount of shared knowledge; it could suggest larger divergence of opinion than was warranted. Therefore, I ran an additional CCM analysis using a binary value to indicate whether two items were together at lower category levels. Because the depth of sorts varied—some individuals subdivided the groups until each wine was by itself—simply examining shared category membership at the terminal nodes of the sorts was inappropriate. Instead, any wines that were together in the bottom third of a given sort were considered together. For example, in a sort with 9 levels, wine pairs with distances of 3, 2, or 1 were coded as 1; all other wine pairs received a 0. The fit of the CCM was comparable to that for the raw distances. Furthermore, correlation of the first three factor scores of these two analyses was high, giving me confidence that the results were comparable and warranting use of the distance results in subsequent analyses.

exploration of potential subgroups via analysis of the first factor scores (the competence scores) and analysis of likely subgroups separately. These analyses are described below.

According to Romney et al. (1986), the first factor score indicates agreement with the group consensus, and therefore, by the logic of the model, may provide an index of expertise. Despite the weak fit of the model, there was a moderate positive correlation between the first factor score and the wine knowledge test (WKT), but no correlation with the tasting and familiarity scores obtained in the card review task (See Table 2). There was a strong positive correlation between familiarity and tasting experience.

Table 2. Pearson correlations among potential indices of expertise

Expertise index	1st factor scores	Tasting	Familiarity	WKT
1st factor scores	1.000			
Tasting	.000	1.000		
Familiarity	.105	.863**	1.000	
WKT	.314*	.090	.161	1.000

** Correlation is significant at the 0.01 level (1-tailed)
* Correlation is significant at the 0.05 level (1-tailed)
Note. The factor loading scores were derived from the principal components analysis of the pair-wise distance values for first sorts only. N = 30 for all cells. The familiarity score is the proportion of the 40 wines that each expert had indicated they were familiar with during the card review task; the tasting score is the proportion they had tasted. WKT refers to the score on the wine knowledge test.

Following Boster and Johnson (1989), I looked for group differences in the factor scores. A series of one-way ANOVAs found no effect of expert type on the first, second, or third factor score. There was not evidence that any of the first, second, or third factor scores reflected the type of expert.

Separately, I looked for evidence of consensus within each of the three expert groups using the CCM on each group's results for the first sort. Running separate analyses for each group obtained only weak fits of the model. Across groups, the ratios of the first to second eigenvalues were small (1.7 to 2.9).

Multidimensional scaling and clustering of sorts

The analyses of the sorting data reported thus far did not reveal the expected similarity among experts of the same type. Overall consensus was weak and did not increase when each expert group was examined separately. To explore whether some factor other than expert type was important, I next conducted MDS analyses of the first sorts from each export to discover what—if any—clusters of sorts might exist. Consistent with the factor analyses, multi-dimensional scaling of the distance matrices did not reveal group-based clusters based on expert type. I used the non-metric MDS ALSCAL procedure (SPSS 16.0.2 for Macintosh, 2008) to create a two-dimensional plot of the experts based on the pair-wise wine distances derived from their first sorts. As described earlier, these distances reflect the number of levels that would need to be crossed to bring a given pair of wines into the same category. The MDS analyses described earlier used these distances to plot the similarity relationships among the wines; in this analysis, the distances allow us to plot the similarity relationships among the experts (treating their first sort as representative). For this dataset, the ALSCAL procedure standardized the variables (by z-score) to compensate for variation in the number of levels per sort (which would affect the absolute distance values) and calculated the Euclidean distances among the experts, assigning them to locations in a two-dimensional space (without rotation).[24] Figure 4 shows the resulting plot of the experts' first sorts. The squared correlation index of the matrix (RSQ = .83) indicates acceptable fit. Although a few clusters of sorts can be discerned, assorted types of experts produced them.

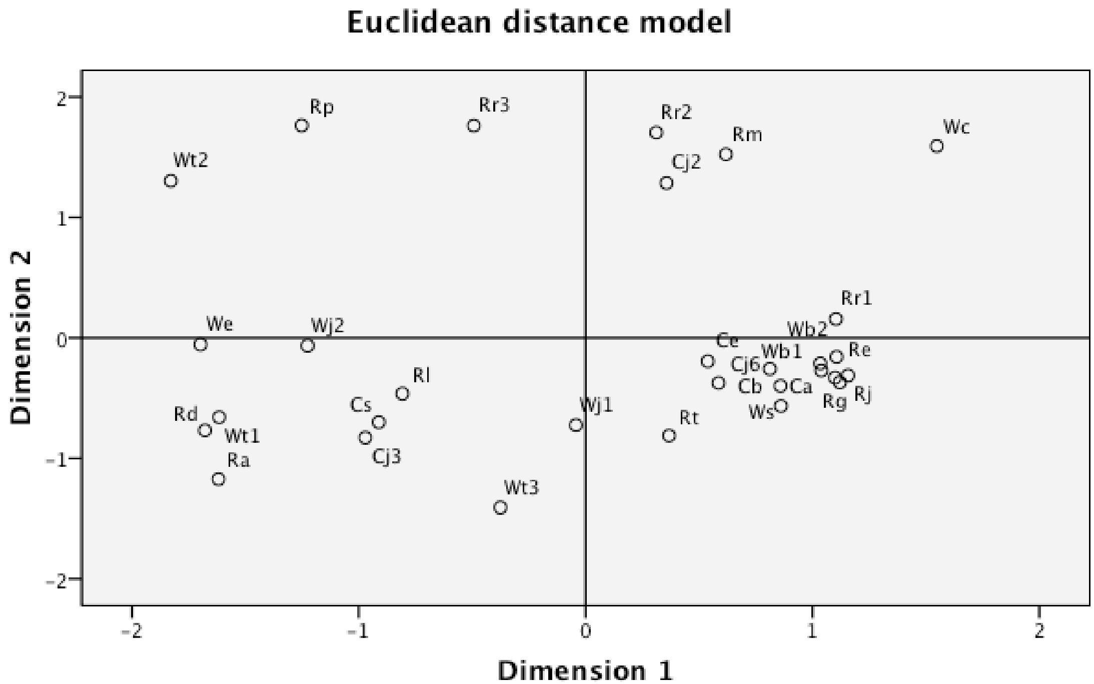


Figure 4. Multidimensional scaling of experts' first sorts

Note: The first letter indicates the expert group: connoisseur (C), retailer (R), or winemaker (W). To obtain this plot, the SPSS ALSCAL program calculated Euclidean distances among experts/sorts based on the 780 pair-wise wine distances derived from each expert's first sort. The distances were standardized using z-scores to correct for variation in the absolute distances.

For the Background Questionnaire, a hierarchical cluster analysis of the activity inventory responses had yielded clear clusters of experts by type. However, the same analysis of

[24] A three-dimensional solution did not add much interpretable information to this analysis (See Appendix L).

sorting distances did not. Although three clusters emerged (marked “A,” “B” and “C” on Figure 5), each combined a mix of expert types.

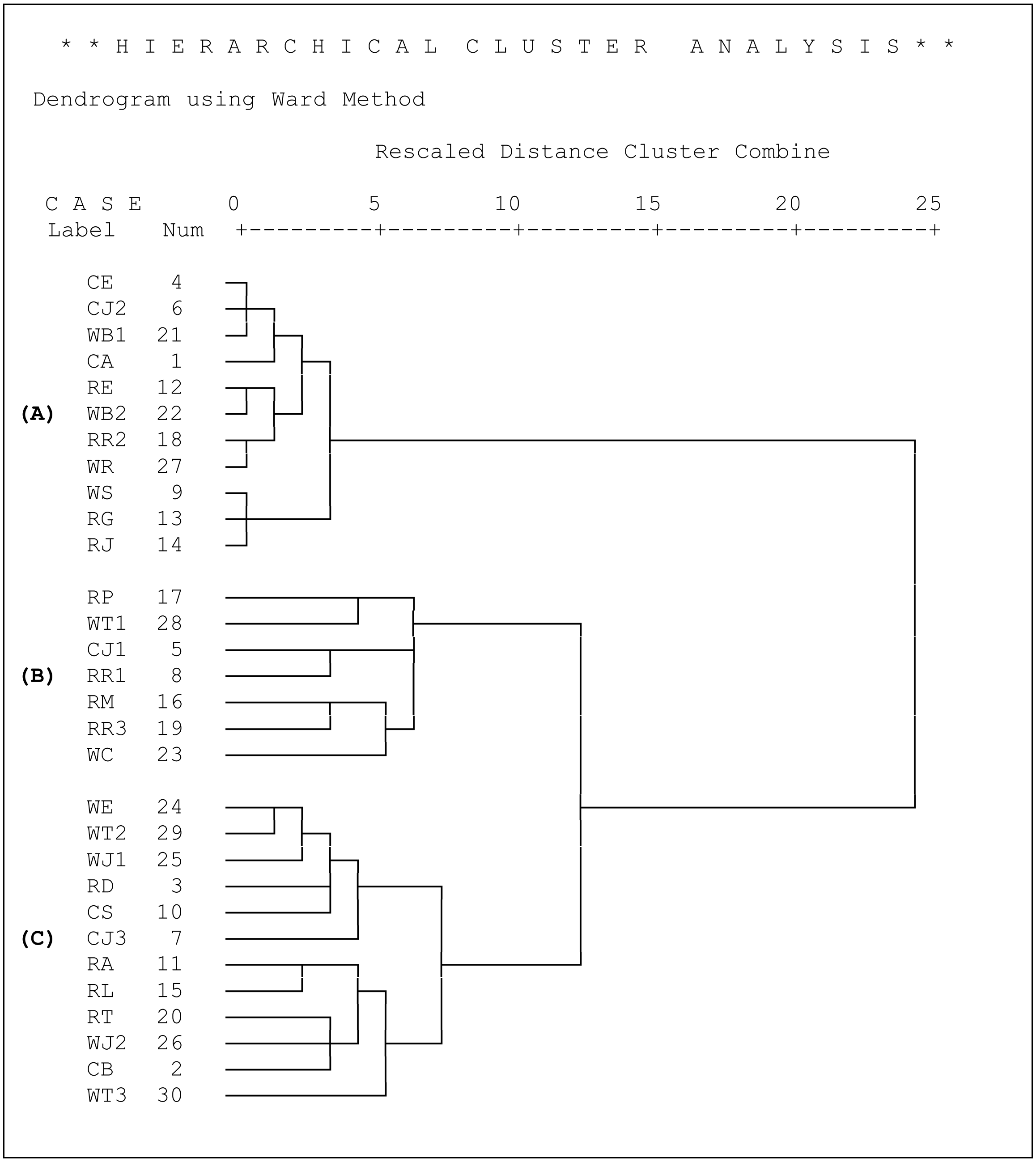


Figure 5. Hierarchical clusters of first sorts

Note: This hierarchical cluster analysis using the Ward method used the pair-wise wine distances derived from experts' first sorts. The first initial of the case label indicates the expert group (connoisseur, retailer, or winemaker).

Superimposing these three groups onto the MDS plot (see Figure 6) reveals that group A maps on to the plot's single tight cluster, which lies midway along the right half of the x axis. Members of group C fall largely in the bottom left quadrant, whereas members of group B are arrayed along the uppermost edge of the plot.

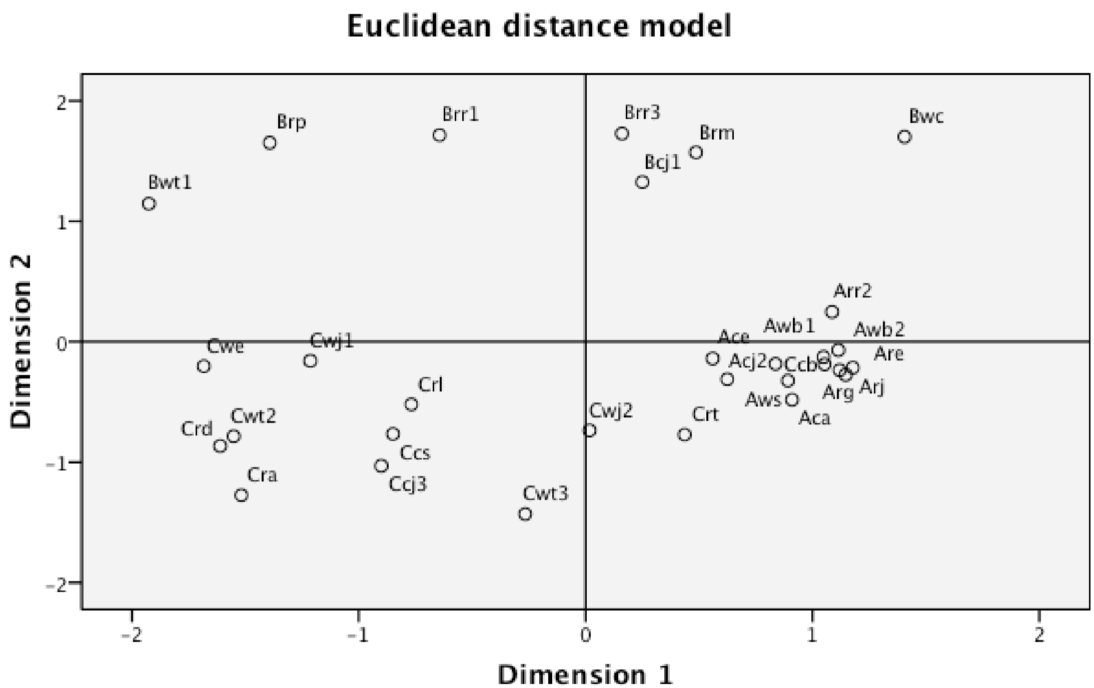


Figure 6. Multidimensional scaling of experts' first sorts, indexed by cluster group

Note: The first letter indicates group based on the hierarchical cluster analysis presented in Figure 5. The second letter indicates the expert group (connoisseur, retailer, or winemakers). Except for the change in labels, this plot is identical to Figure 4.

Analysis of the structural characteristics revealed the following differences. A one way ANOVA on the number of levels found a significant effect of cluster, $F(2,27) = 6.94$, $p = .00$, with type A sorts having more levels than type B or C sorts (ps < .05 with Bonferroni adjustments). See

Table 3 for descriptive statistics. A separate ANOVA also revealed a significant effect of cluster on the number of categories at the initial level, $F(2,27) = 9.03$, $p = .00$, with type C sorts having more categories initially than either of the other two sort types (ps $< .05$ with Bonferroni adjustments). There was also a significant effect of cluster on the average number of categories, $F(2,27) = 3.99$, $p = .03$, with type C having a higher average than type B (ps $< .05$ with Bonferroni adjustments). In sum, type A sorts were deepest, type B sorts were narrowest and type C sorts were most broad. In the next section, I supplement these descriptions with information about the types of category labels that characterized the three sort types.

Table 3. Structural characteristics of sort types

		Sort type		
		A (n = 11)	B (n = 7)	C (n = 12)
Levels	Mean	4.45	2.86	2.83
	SD	*1.57*	*.69*	*.83*
Categories at initial level	Mean	8.00	8.00	20.83
	SD	*3.85*	*5.42*	*11.48*
Average categories per level	Mean	10.31	8.57	14.08
	SD	*3.44*	*3.79*	*5.41*

Summary of similarity analyses of sorting data

Although the expected expert groups had emerged from analyses of the behavioral reports collected in the Background Questionnaire, these same groups did not emerge from the sorting data. Multidimensional scaling and cluster analyses suggested some groups among the sorts, but these did not correspond to the three expert types. The cultural consensus model was only a weak fit to the sorting data, suggesting that the consensus among the experts was not strong. Analyses of the similarity among wines suggested that color and region were important dimensions. In the next section, I look more closely at the category labels experts used.

Content

What principles did experts use to organize the set of wines? I coded the category labels for content to identify the most common types of labels and explore group pattern of category use. The behavioral data had led me to expect some group differences in category use. For example, winemakers might focus more on issues related to wine production and less on regional categories. Connoisseurs might emphasize evaluation of wines. However, the absence of group-based consensus for the sort distances weakens the strength of these expectations. At the same time, analysis of sort content may help explain what is driving the three mixed-expert clusters of sorts (A, B, and C) that emerged from the hierarchical cluster analysis and multidimensional scaling.

Category coding

To analyze the content of experts' category labels I coded them using the following eleven category codes: (1) color, (2) combination, (3) grape, (4) preference, (5) price/quality, (6) process, (7) region, (8) role, (9) style, (10) type (sparkling/still), and (11) other.[25] The color code referred almost exclusively to two category names ("Red" and "White"). The combination code

---

[25] The category codes were developed while reviewing an alphabetized list of the 637 unique category names that experts gave for the groups of wines they created during the sorting process. I created codes to try to describe the content of the name. Most categories were assigned sub-types (e.g. continent, country, state, and appellation were subtypes of the region category), but these have been omitted from the analyses here. After coding all categories and tallying the responses, I consolidated the less frequent types (those representing fewer than 5% of any group's categories). Many of these were lumped into the *other* category, but it seemed appropriate to combine others with existing groups (e.g., appellation joined *region* and distinct production categories were all grouped under *process*).

described categories mixing two or more category types (e.g., “Dry reds” or “Italian sparklers”), but also applied to “polysemous” labels (labels that were difficult to interpret because the terms had multiple potential referents). For example, “Champagne” could be interpreted as referring to region, type (sparkling/still), process (“methode champenois”), and/or a particular blend of grapes. In coding the combination categories, I specified the basic category types that were being combined. So, for example, the category name “Dry white Italians” was coded as a combination category type with a style/color/region subtype. The grape category type referred mostly to varietal names (e.g., “Merlot”), but also to broader categories such as “Hybrids” or “Vinifera grapes.” Preference categories were organized around personal interest or liking (e.g., “Wines I’d like to try”). Price and quality were combined because price-based categories were relatively rare and it was sometimes hard to differentiate the two category types (e.g., “High end wines”). Process categories described wine production in terms of a variety of issues such as fermentation technique (“Botrycized”), the control of grapes (“Winery-grown fruit”), the composition of the wine (“Single varietals”) and occasionally even the specific winemaker (“A Gaja wine”). Region often referred to country names, but the specificity of the categories ranged from broad categories (e.g., “Old World wines” or “Imports”) to extremely narrow ones (e.g., “Mendocino appellation”). Role categories referred to the wine-drinking situation (e.g., “Dessert”). Style categories grouped wines in terms of their flavor, weight, and other taste-oriented characteristics; “Sweet” and “Dry” were two very common style categories. Type (sparkling/still) categories distinguished sparkling from still wines. Finally, the *other* category type captured responses that were vague or difficult to interpret as well as less common types.

Because of the variability in the number of categories generated by the three groups, the frequency information in Table 4 is presented as the percentage of each group's categories that received the designated category type code (and as a percentage of all categories for the Grand Total column). Separate one-way ANOVAs for each category type found a significant effect of expert type for Grape ($F(2, 67) = 2.21$, $p < .05$), with Tukey post-hoc tests showing that connoisseurs used more grape-based categories than retailers ($p < .05$). There were marginally significant differences between winemakers and retailers for grape and preference.

Table 4. Category types as a percent of each expert group's categories

Category type	Connoisseurs	Retailers	Winemakers	Grand Total
Color	14.03%	10.09%	6.79%	9.83%
Combination	20.52%	21.25%	22.75%	21.62%
Grape	16.88%	5.05%	12.56%	10.57%
Other	4.94%	7.65%	7.98%	7.13%
Preference	0.52%	0.76%	6.96%	2.95%
Process	2.08%	2.75%	5.43%	3.56%
Region	24.16%	33.03%	24.96%	28.01%
Style	1.82%	7.19%	4.07%	4.79%
Type	4.68%	6.12%	2.55%	4.48%
Price/Quality[26]	10.39%	6.12%	5.94%	7.06%

[26] It is notable that only one of the connoisseurs' 40 price/quality category names is based on price. For retailers and winemakers it was a near even split between the two (the ratios of price to quality categories were 21:19 and 18:17, respectively).

The most common category type was region (28%), followed closely by combination (22%). I'd like to focus for a moment on these combination categories. These figures may underestimate their importance, because while 92% of the sorts mixed more than one category type,[27] only categories whose labels explicitly mixed category types (or were polysemous), were coded as combinations. In reality many of the groups of wines were implicit combinations of multiple category types (e.g., a regional group that had been subdivided into varietal groups[28]). Explicit combination category labels were used at a variety of sort levels, but were most common at the initial level of sorts (43% of the 352 explicit combination categories).

Because combination categories constituted a substantial proportion of categories I examined the contribution of the other nine category types via the combination categories. To do so, I coded each component of a combination (i.e., each subtype) separately. Thus, "Dry white Italians" would be assigned three separate codes instead of one. Subsequent analyses use these "decomposed" category codes.

Category type use by individual and by expert group

To what extent did individuals from each group use different types of categories? Appendix M presents the proportions of experts from each group who used each category type at least one time. Chi-squared analysis of these frequencies did not reveal significant differences in the three groups' use of the categories for all sorts or for the first sorts.

---

[27] Of the six "pure" sorts, two were second sorts and four were third sorts. These sorts used less common category types: price (three sorts), quality (two sorts), and "other" (one sort focused on the effectiveness of the labels). The sorts were shallow, with either a single level (four sorts) or two levels (two sorts) and had relatively few total categories (range = 2 to 7 categories per sort).

[28] These intersections were not coded as combination category types. Even if the expert used the "parent" category's name when giving a label, I removed the "pedigree" information from the label for the coding process.

The absence of a significant difference in the numbers of individuals using the category type suggests they had equal access to the different types of categories, yet they may have employed them to different degrees. Exactly how to test this hypothesis is a little tricky. Using raw category counts is problematic because the individual variation in the number of categories was so large. Perhaps more serious is the fact that in a multi-level sort that uses a mix of category types, the point at which a person decides to use a particular strategy may have dramatic effects on the counts for different category types. Take, for example, two subjects who each create a single two-level sort. Both of them use two category types: color and region, breaking color into two categories and region into five. If Person A starts by grouping the wines by color and subdivides by region, that person will have two color categories and ten region categories. If Person B reverses the procedure, subdividing the regional categories by color, that person will have five region categories and ten color categories. Using the counts (or averages) makes it look as if region is more important to Person A than B, when that is probably not the case.[29] To address this issue, I calculated the proportions of each category type by level and averaged those for each participant. Figure 7 presents the means by category type averaged across all experts; Figure 8 breaks them down by expert group.

A 3 x 9 ANOVA (Group by Category type) on the category type proportions per level found a significant effect of category type, $F(4.3, 115.0) = 17.21$, p =.00 and a marginal category type by expert group interaction, $F(8.5, 115.0) = 1.78$, p = 08.[30] Pair-wise comparisons with Bonferroni adjustments to alpha levels revealed a number of significant differences in the

[29] One might even argue the reverse—that region was more important to person B, because that was the first type of category division imposed.
[30] Because the assumption of sphericity was not met, I applied Greenhouse-Geisser corrections, which adjusted the degrees of freedom from (8, 216) and (16, 216) to those reported.

relative use of the nine category types. Region (M = .28, SD = .17) was again the dominant category type, used more often than style (M = .09, SD = .10), $t(29)$ = 4.74, p = .000, grape (M = .08, SD = .08), $t(29)$ = 5.37, p = .000, process (M = .06, SD = .11) , t (29) = 5.46, p = .000, preference (M = .05, SD = .08), $t(29)$ = 5.89, p = .000, type (M = .05, SD = .04), $t(29)$ = 7.28, p = .000, and role (M = .03, SD = .05), $t(29)$ = 7.74, p = .000. Color (M = .16, SD = .12) was used more often than preference, $t(29)$ = 4.11, p = .00, process $t(29)$ = 3.27, p = .003, type, $t(29)$ = 4.71, p = .000, and role, $t(29)$ = 5.60, p = .000. Price/quality (M = .12, SD = .11) and grape were used more often than role, $t(29)$ = 4.01, p = .000 and $t(29)$ = -3.22, p = .003, respectively.

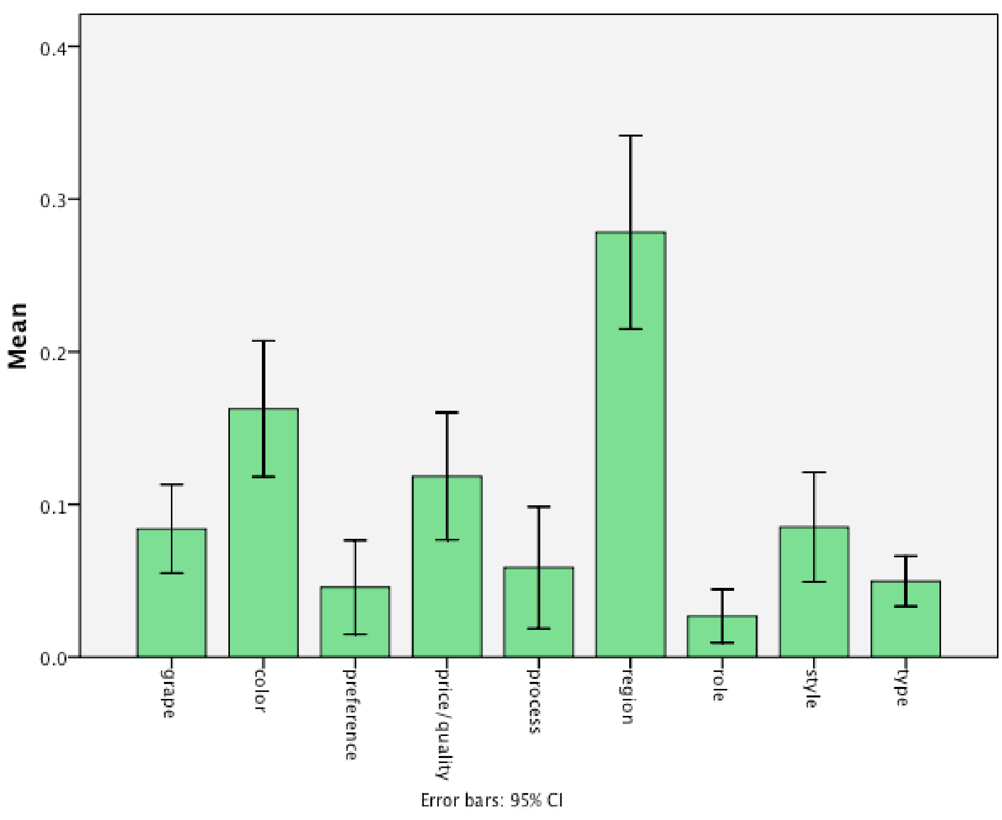


Figure 7. Mean category proportions per level (averaged across all sorts for each expert)

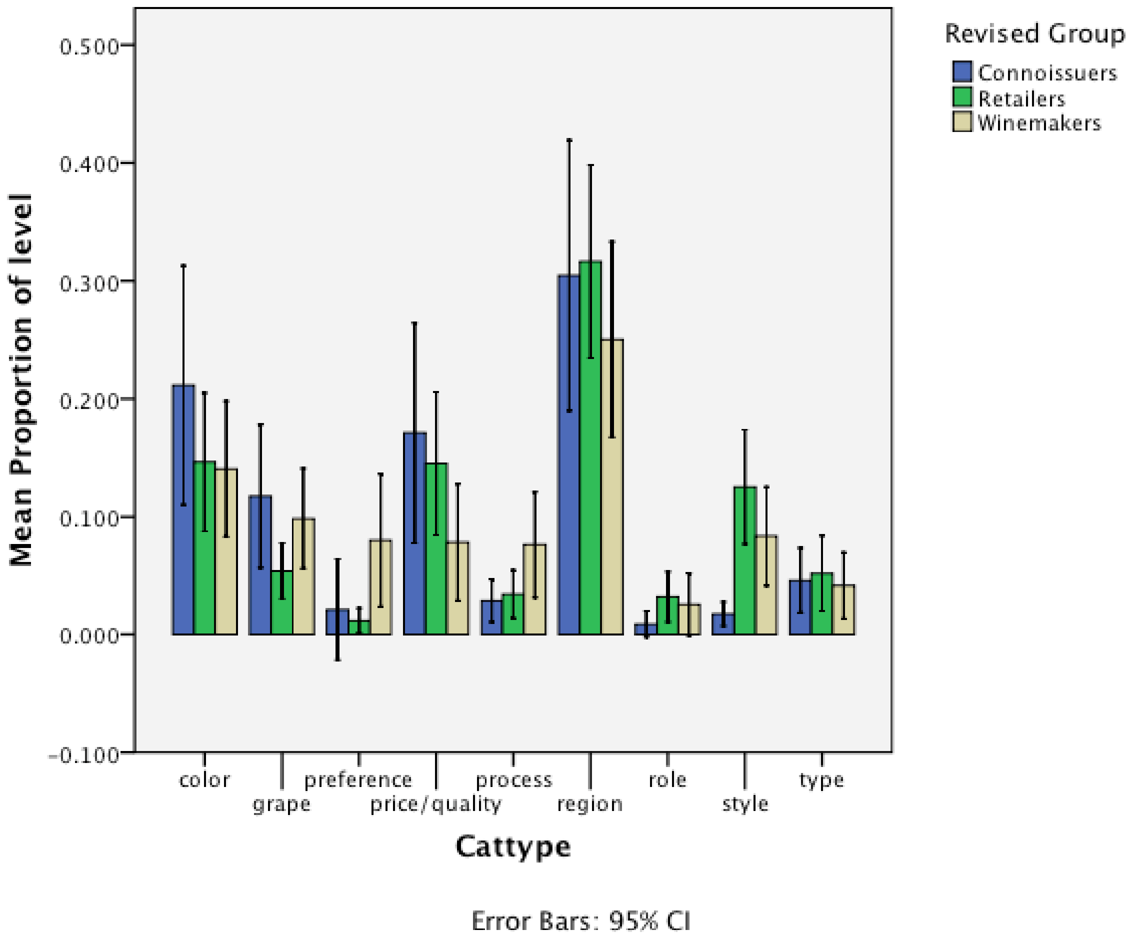


Figure 8. Mean category type proportions by expert group (all sorts averaged by expert)

Note: This figure and the analysis of the impact of group type on mean proportion per level used all sorts. That is, the proportion of each category type for each level of each expert's sort was averaged for that expert and then averaged for the group. This kept each expert's contribution to the group mean comparable. The results were not very different when only the proportions (by level) from the first sort were included in the group means.

Post-hoc examination of the three expert groups' reliance on the different category types separately identified only two category types for which there was a significant effect. Retailers used style-based categories more than connoisseurs, $F(2, 27) = 4.36$, $p < .05$, with Bonferroni adjustments. Winemakers relied on preference-based categories more than retailers, $F(2, 27) = 4.139$, $p < .05$.

Note that while there were few statistically significant differences among the groups' use of specific category types, there are several additional category types where the means are consistent with the predictions. Within group variation, however, swamps the effect. In particular, there seems to be a trade-off between preference and price/quality, with winemakers' significantly higher mean on preference being compensated for by a (not significantly) lower mean on price/quality. As expected, winemakers also have a lower mean for region than the other two groups, and a higher mean for process. One possible source of the within group variation would be the availability of multiple conceptual organizations. To the extent that winemakers, for example, have and use alternatives to a process-based categorization strategy, that will make use of process appear inconsistent.

Category type use by sort cluster

Given the absence of strong group-based differences in the similarity data, it is not surprising that there were not large group-based differences in the content of the sorts. Perhaps expert type was not the critical variable. Remember that the hierarchical cluster analysis identified three groups of sorts that cut across occupational lines. What underlies the clusters, if not expert type? I next examined the content and structure of the sorts in those three groups (A, B, and C), to try to identify interpretable regularities.

I examined the proportional use of category type by level, as above. To simplify the presentation of these results (Table 5), I have omitted those levels that were used by only one or two experts in the cluster[31] and have also collapsed several category types.

Table 5. Proportions of category types by level for sort types

Cluster set	Level	color, grape	preference, price/quality	region	other, process, role, style, type
A	Up 1	0.00	0.00	0.91	0.09
	Initial	0.02	0.00	0.95	0.02
	Down 1	0.35	0.07	0.26	0.32
	Down 2	0.24	0.22	0.27	0.27
	Down 3	0.20	0.02	0.43	0.36
A Total		0.23	0.09	0.42	0.26
B	Up 1	0.00	0.71	0.00	0.29
	Initial	0.11	0.50	0.13	0.27
	Down 1	0.33	0.12	0.11	0.44
B Total		0.21	0.25	0.19	0.34
C	Up 1	0.52	0.04	0.15	0.30
	Initial	0.36	0.09	0.32	0.23
	Down 1	0.28	0.24	0.23	0.25
C Total		0.33	0.16	0.26	0.26

[31] For group A five levels were omitted, four below the initial level and one above. For group B, one level below the initial level was omitted. For group C, four levels were omitted, one above the initial level and three below it.

Examining the proportions of different category types by sort levels revealed three general profiles that I've characterized as follows: (A) Regional: At the initial level these sorts almost exclusively used region-based categories and continued to emphasize region throughout a deep sort (of at least five levels). (B) Evaluative: These were relatively shallow sorts (typically three levels) that depended heavily on evaluative categories (price/quality or preference) at the initial level and above. (C) Color/Grape Combination: The majority of the initial level categories for these sorts were combination categories that emphasized color and grape as well as region.

Discussion of content analyses

Can the differences in the three expert groups' use of different category types be explained in terms of their distinctive goals and behaviors? It would make sense for winemakers to be the group most interested in process. For retailers, focus on the style of wines may be important for effective customer communication. The two evaluative categories (preference and price/quality) are used differentially, with winemakers grouping wines more by personal taste and connoisseurs grouping them by "objective" quality considerations. This is interesting, given that connoisseurs are engaged in a relatively personal endeavor (the enjoyment of wine) whereas winemakers are presumably trying to meet more "objective" standards. However, it is also true that winemakers (and retailers) are in the business of pleasing people with quite diverse tastes—what one person enjoys another may not. This may make them more sensitive to the subjectivity of wine preference. Connoisseurs, in contrast, may be more focused on their own taste and developing that with reference to the judgments of established authorities. Remember that their only distinctive activity on the Background Questionnaire inventory was the study of wine rankings and ratings.

Discussion

The results of the sorting task clearly demonstrated that wine experts have access to multiple ways of organizing the domain. Not only did they easily generate multiple sorts, but also within these sorts they integrated a variety of dimensions to a surprising degree. Even at the level of the individual category, the experts frequently mixed dimensions, making the "combination" category type the second most common overall.

Analysis of the category labels revealed that the most commonly used "pure" categories were those based on region and color; the same dimensions that emerged as important in the multidimensional scaling solutions of the experts' sorting distances.

However, factor analytic techniques (CCM, Romney et al., 1986) revealed neither strong overall consensus nor consensus among subgroups defined by expert type. Neither were there many differences in the types of category labels used by the three expert types. The only reliable group differences were that winemakers were more likely than other experts to use category labels emphasizing preference whereas retailers were more likely than the others to use style-based categories. There appeared to be substantial variation within the expert groups. Hierarchical cluster analyses of the experts' first sorts produced three groups that mixed the three expert types. However, closer examination of the sort clusters that did emerge revealed some consistency in the types of categories they emphasized and the structure of the sorts. One cluster (type A) tended to be deep and emphasize regional distinctions. A second set (type B) was shallow and used evaluative types of categories. The third large cluster (type C) used more combination categories, especially integrating color, grape, and region category types. The three sets were not equally well defined; multi-dimensional scaling of the sorts found tighter clustering of type A, with the other two groups being more dispersed, though discernible.

CHAPTER FOUR: CATEGORY TESTING (SESSION TWO)

Overview

In the second session, I evaluated the psychological "reality" of the categories generated in the first session. Specifically, similarity judgments, category membership ratings, and category ratings were used to assess the availability and importance of the categories for the different groups. In addition, an inference task tested the relative utility of the categories. Before delving into the specifics of each task, I will describe procedural details common to all Second Session tasks.

Participants

The same 30 wine experts who participated in the first session were contacted to complete a second session. Of these 10 retailers, 9 winemakers and 5 connoisseurs completed the session. All expert interviews were conducted within a seven-week timeframe. The 11 novices completed the similarity task only, and did so at the same time that they completed the Wine Knowledge Test and abbreviated Background Questionnaire described in Chapter 2.

Protocol

All participants completed the same set of tasks in the same order, as follows: (1) consent form review, (2) silent review of wine label cards, (3) similarity judgment task, (4) category membership rating, (5) inference task, (6) review of acknowledgment form and debriefing.

In subsequent sections I discuss the materials, procedures, and results for Steps 3, 4, and 5 individually. Step 1 used a consent form identical to the one from the first session. In Step 2, participants quickly reviewed the set of 40 wine label cards used in the first session to refresh their memories. In Step 6, they completed a form (see Appendix N) indicating whether they wanted to be thanked by name in the acknowledgments. At that time I thanked and debriefed participants, commented on preliminary results and gave them a copy of the debriefing sheet to keep (Appendix O). The sessions varied in length from approximately 45 to 150 minutes, but typically lasted about an hour.

Objectives

The overarching goal of Session 2 was to assess the psychological reality of the categories that had been generated in Session 1. Task demands may have influenced the output of the sorting task; therefore the initial categories must be interpreted with some caution (Ross & Murphy, 1999). It is possible that they do not reflect participants' true conceptual organization. Thus, the aim of the second session was to examine a subset of categories from a variety of angles, using the following measures.

The similarity judgments (Step 3) were designed to assess the availability of different types of categories. The category membership ratings (Step 4) provided an indication of the extent to which the specific categories examined were key components of the wines' feature structures. If a single wine was considered an excellent member of multiple categories, this provided evidence of cross-classification. The most important task was the final one (Step 5), which examined which categories were used in inductive inference. The key question was whether the experts preferentially used the types of categories they had generated more often in the sorts, or whether the subtle differences evidenced in the sorting task disappeared in the context of a specific problem. For example, in the sorting task, Group A may have emphasized category type X while Group B emphasized category type Y. The inference task explored whether the groups remained "loyal" to their preferred category type (i.e., A sticks with X and B sticks with Y) or, whether both groups had access to both category types and deployed them as needed (i.e., both A & B use X for inference tasks of type x; both A & B use Y for inference tasks of type y).

Similarity

The psychological "reality" of the categories that emerged from the sorting tasks was suspect because of the potential influence of task demands. The goal of the similarity judgments task was to obtain confirmatory evidence that the sorting categories were available to experts and to make comparisons across expert types.

In this task, participants evaluated the overall similarity of two wines. For half of the trials a category label was presented, for the other half it was not. This design is an adaptation of Barsalou's test of context-dependence and context-independence (Barsalou, 1982; Ross & Murphy, 1999).

According to the logic behind the task, similarity judgments depend on the sets of features being evaluated. Barsalou (1982) distinguished between those features that are always available (context independent) and those that depend on local information (context dependent). To the extent that priming a category name facilitates the retrieval of features that lead to higher similarity judgments, those features are context dependent. If the presence/absence of the prime has no effect on similarity judgments, the features that the items share must be category independent (Barsalou, 1982; Ross & Murphy, 1999). By corollary, categories that require priming to exert their influence must be less available than those that do not.

I used this technique to assess the relative availability of different types of categories. To the extent that the presence of a label increases similarity relative to no label, this suggests that the category is context dependent. I planned to interpret such positive shifts in similarity as evidence that the category was less "real" to the experts than were the categories for which the presence of a label did *not* increase similarity. As it turns out, the experiment did not work in exactly the way I expected. I will detail the "surprise" in the results section.

One concern about the validity of this interpretation (Rips, personal communication, January 2007) is that no change in the similarity rating could be an indication that the labels were irrelevant, not support for their availability. To address this, I sought frequently generated categories and wines that at least some experts had listed as members of those categories. In addition, in Task 5 participants rated each wine's membership in a subset of these categories.[32] Analysis of these results should address concerns that the labels were irrelevant.

Methods

Design

The similarity task had a three-factor, mixed within- and between-subjects design. Category type and label presence/absence were within-subjects factors. Expert type was a between-subjects factor.

I chose 13 categories from the more than 600 category names generated in the sorting task. For each of these 13 categories, I selected two pairs of wines. The specifics of how I selected the categories and the wines is described below under the heading "Selection Process." One wine pair from each category was randomly assigned to "Set A" and the other to "Set B." Each subject evaluated the similarity of all 26 wines pairs, viewing one set without labels (LA = label absent) and the other set with labels (LP = label present). For half of the experts, Set A did not have labels and Set B did; for the other half, Set B did not have labels and Set A did.

---

[32] Time constraints prevented assessment of all the categories from this task.

All 13 comparisons for a given condition were blocked together and recorded on a single sheet of paper. Each of the four set-condition combinations used a unique, randomly assigned order of the 13 trials. The four block orders used were (1) ALA-BLP, (2) BLA-ALP, (3) BLP-ALA, and (4) ALP-BLA.[33]

Materials and participants

Participants included eleven novices as well as the twenty-four experts who completed all three of the second session tasks. For each trial, participants viewed reduced, black-and-white copies of two wine labels printed side by side on an 8.5 x 11 sheet, landscape orientation. The question, "How similar are these two wines?" appeared below the wines for the label absent condition. For the label present condition, the name of the category appeared above the images and the question, "How similar are these two <category name> wines?" appeared below the wines. See Appendix P for examples.

For each block, participants received a rating sheet listing the 13 wine pairs. The pairs were represented on the rating sheet by two numbers, which were also printed on the comparison pages. A scale for recording their similarity judgment as a number from 1-7 appeared to the right of the pair. See Appendix Q for a sample rating sheet.

Procedure

Participants heard the following instructions as they received the similarity rating sheet and a notebook containing the comparison pages for their review.

> I am interested in what you think about the similarity of different wines. On the following pages you will see pairs of wine labels.

---

[33]A combination of factors including participant attrition, the reassignment of connoisseurs to other expert groups, and the late addition of blocks 3 and 4 resulted in an unbalanced treatment for the remaining connoisseurs.

> Please think about the wines represented by the labels and rate how similar they are using the 1 to 7 scale at the top of this score-sheet. Circle "1" on the line to the right of the pair if you think the wines are not at all similar and "7" if you think they are very similar. Use the numbers between 1 and 7 as appropriate. Remember that I am asking you to judge the similarity of the wines, not the labels.

The participants then worked through the 13 pages in the set, recording their responses on the rating sheet. After participants had completed the first block, they were handed the second rating sheet and heard the following additional instructions.

> Now you're going to repeat the process with a new set of wines. This time, you'll see the name of a wine category, followed by two wine labels. Again, judge the overall similarity of the two wines, recording your answer using the 7-point scale. On each page, please be sure to read the label and the question at the bottom before responding.

Note that the instruction texts above were used when the Label Absent block preceded the Label Present block. When the order was reversed, the instructions were adjusted appropriately.

Rationale for Groupings

Murphy and Ross's (1999) research on food categories, served as my template for this task. They used it to compare the availability of taxonomic categories relative to the script-based ones that had emerged in their sorting and category generation tasks. For the similarity items I considered three broadly defined groups of categories: "taxonomic," "script," and "combination."

Taxonomic Categories

Analysis of the categories that emerged from the Session 1 sorting task showed that Region, Grape, and Color were the predominant category types across all subject groups, with some differences in emphasis among the groups.

The Grape category type refers to the use of varietal categories and thus echoes the use of folkbiological categories in other research. However, while varietal may be part of a folkbiological taxonomy, where wine is concerned, the primary focus is at the very lowest level of the taxonomy: the question is which varietals of the *vitis vinifera* species are ingredients. Most exceptions to this generalization are North American wines that use *vitis labrusca* or hybrids. Because varietals are all members of the same superordinate category, the structure is broad and shallow.

One might consider the Color category type to be a superordinate of varietal, because certain varietals tend to be used to create the same color of wine. For example, Sauvignon Blanc grapes usually yield white wines whereas Syrahs usually yield reds. Technically, however, the ultimate color of a wine depends on the treatment of the grape skins, and is not due exclusively to the type of grape pressed. (It is rare, however, to see the same varietal used to make both a red and a white wine.) Further complicating the issue, many wines are blends of multiple varietals. For example, traditional champagnes are blends of Chardonnay (typically a white), Pinot Noir (typically a red), and Pinot Meunier (typically a red).

The Region category type is also "sort of" taxonomic. One can certainly group regions in a hierarchy. Country (one of the most commonly generated types of categories in the sorting task) is an intermediate level category that can be both lumped and split. Countries are lumped into superordinates such as continent or other larger regional groupings (e.g., "new world"). Countries can also be split up into smaller categories such as states or other within-country regions. These smaller categories can be further subdivided into AVAs or appellations, or further still to the level of specific vineyards.

Yet despite this structural appropriateness, there are problems with treating regional categories as taxonomic for wine. Taxonomic categories typically describe sets of objects that share many common features, but there is a great deal of variation among wines from a single region. Although many wine experts emphasize the importance of terroir, the ultimate characteristics of a given wine do not depend on terroir alone, but are the result of a particular combination of factors including the terroir, the grape(s), the winemaker's technical choices, and the weather over a specific time period. In addition, Region is less about the physical "matter" of the wine and more about its historical context, something that has some strong script-like overtones. Despite reasonable objections to calling Grape, Color, and Region taxonomic categories, they are the strongest candidates.

I wanted to include categories at a variety of levels of specificity, so I used five "taxonomic" categories.

<u>Taxonomic categories used in similarity task:</u>

Region-Superordinate (Super-country)	"Old World wines"
Region-Basic (Country)	"Italian wines"
Region-Subordinate (State/Sub-country)	"California wines"
Grape-Superordinate (Color)	"Red wines"
Grape-Basic (Varietal)	"Sauvignon Blanc wines"

As described earlier, I used two pairs of wines from each category. I'll discuss the selection of those wines shortly.

Script Categories

Ross and Murphy (1999) observed many script categories in their research in foods, and I also found some.[34] The script categories were diverse and not obviously divisible into more and less abstract levels of specificity, though general themes of production and purchase seemed to emerge. The production scripts addressed winemaking techniques, wine composition, or other aspects of the production process. The purchase scripts addressed who bought the wine and for what purpose. Note that even with such broad groupings, it was not always easy to disentangle production and purchase. I included five script-like categories and tried to vary the themes:

Script-ish categories used in similarity task:	
Purchase (Situation)	"Party wines"
Production/Purchase (Taste)	"Sweet wines"
Production/Purchase (Quality/Price)	"High End wines"
Production (Technique/Taste)	"Sparkling wines"
Production (Composition)	"Single Varietal wines"

Combination Categories

In addition to the taxonomic- and script- categories, many categories combined category types. For example, the category "California reds" (region/color) explicitly combines two taxonomic categories. Yet even when categories were not explicitly labeled as a combination, they sometimes constituted implicit combinations. Because of the general tendency among

[34] To reiterate, however, the boundary between the taxonomic and the script categories is ambiguous—I am not certain this distinction would have jumped out at me had I not been looking for it.

experts to mix and match the type of category across levels (e.g., they sorted by Region initially but divided the second sort by Grape), many categories that were explicitly taxonomic were implicitly combinations.

Although combination categories were common, I used only three in the similarity task, for several reasons. First, the explicitly labeled combination categories were generally quite small, including only a few wines in any given one. I needed at least three wines per category to generate the two pairs needed for the experiment. Second, although combination categories were a common type, not many of the specific combination categories (e.g., 'French chardonnays") were very common. In choosing combination categories for this task, I included one combination of the two taxonomic types with each other (region/grape) and one of each of them with a script category (region/script and grape/script).

Combination categories used in the similarity task:

Region (Sub) and Script (Production)	"Bordeaux blends"
Region (Sub) and Grape (Basic)	"California Chardonnays"
Script (Taste) and Grape (Super)	"Light whites"

Selection Process

Categories.

In choosing which categories to use, I tried to identify ones that were fairly common overall (i.e., that had high counts in terms of the raw numbers of categories that had taken the label). In addition, I gave preference to those that had been used by multiple experts, preferably different types of experts. Identifying the first few that met the criteria was easy, but became more challenging as I went on and had to dig deeper.

As mentioned earlier, in order to generate two pairs of wines for each category, the set needed to include at least three wines that were members of that category. Using a subset of expert sorts, I created a database to identify the specific wines that were often included in a number of common categories. This database allowed me to focus my attention on those categories that had at least three (relatively consistent) members. It also provided a menu of wines to select from in choosing the wine pairs, as outlined below.

Wine Pairs.

Ross & Murphy (1999) had used independently obtained typicality ratings to try to ensure that the items included in their similarity pairs were of approximately equal typicality. I did not have that data available when designing the task. In selecting items, I relied instead on the sorting task for a rough, indirect measure of typicality. To choose the pairs, I focused on wines that were identified as members of the category by multiple experts (trying to avoid wines that had been included erroneously).

In their design, Ross & Murphy (1999) also made efforts to ensure that only the category of interest was driving the similarity judgments. To do this, they sought pairs of foods that did not share membership in other categories. This was fairly straightforward for their stimuli, because their script and taxonomic categories were mostly independent of each other (i.e., grains are served at a variety of meals, so it was easy to pair a breakfast grain with a dinner grain). I attempted to follow the same guidelines. That is, when I considered the list of wines included in a category, I tried to choose pairs that did not have much else in common by reviewing an informal checklist of dimensions (e.g., color, grape, region, price/quality, role, style). Thus, for the "Italian wines" category I created pairs that were different in terms of color, grape, price, role, and style. In other words, I tried to separate separable dimensions.

However, it was not always possible to vary all of these factors for two reasons. First, the set of wines I had to choose from was limited. Second some categories were (near) subsets of more abstract ones. For instance, all of the "Bordeaux blend wines" were "Red wines." This means that more-abstract categories were able to include less-similar wines, whereas the wines selected for the more-specific categories necessarily had more in common with each other.

Results

Did the presentation of the category label affect similarity judgments? Averaged across all participants, there was no effect of adding a label (see Figure 9).

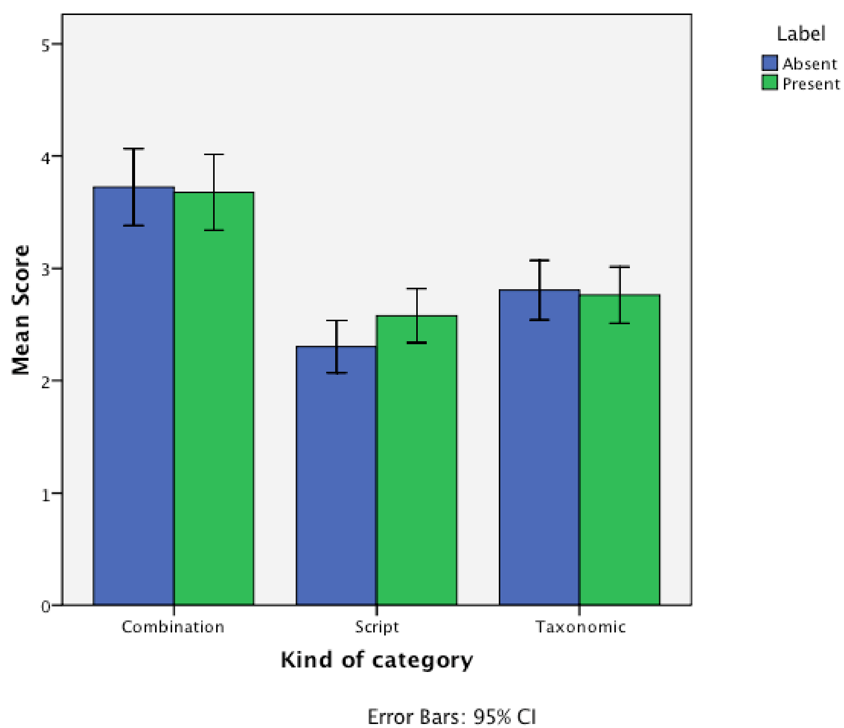


Figure 9. Average similarity score by label condition and kind of category

There is an effect of category type, with pairs of wines from the same combination category generally receiving higher similarity ratings. The absence of any effect of label was unexpected, given the results of comparable experiments (Ross and Murphy, 1999; Barsalou, 1982; Nguyen, 2007). These prior studies had all found similarity judgments to increase when a label is present. To determine whether the difference was due to the expert status of some participants I ran additional analyses.

To examine the impact of expertise, I conducted a three-way ANOVA on the between subjects factor of expert group (connoisseur, novice, retailer, and winemaker) and the two within-subjects factors of label (present, absent) and category type (combination, script, and taxonomic). See Tables 6 and 7 below. This analysis found significant effects of category type, $F(2, 64) = 53.12$, $p = .00$, and of expert group, $F(2, 32) = 3.91$, $p = .02$, but no significance for label. There was, however, a marginally significant interaction between label and expert group, $F(3,32) = 2.53$, $p = .07$.

Post-hoc tests established that novices gave higher similarity ratings than retailers (LSD, $p = .002$). For the main effect of category kind, combination pairs received higher similarity ratings than taxonomic pairs, which received higher ratings than script pairs (with Bonferroni adjustments, *p*s = .00).

Table 6. Mean similarity ratings by group

			95% Confidence Interval	
Group	Mean	Std. Error	Lower Bound	Upper Bound
Connoisseurs	2.81	.28	2.24	3.37
Novices	3.40	.18	3.03	3.77
Retailers	2.48	.21	2.05	2.90
Winemakers	2.99	.20	2.59	3.340

Table 7. Mean similarity ratings by kind of category

Kind	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Combination	3.65	.14	3.36	3.95
Script	2.37	.13	2.10	2.64
Taxonomic	2.74	.12	2.50	2.98

It appears that while presence of a label has little effect on the similarity judgments for experts, further examination with a larger number of subjects might show that adding a label increases perceived similarity for novices (see Figure 10).

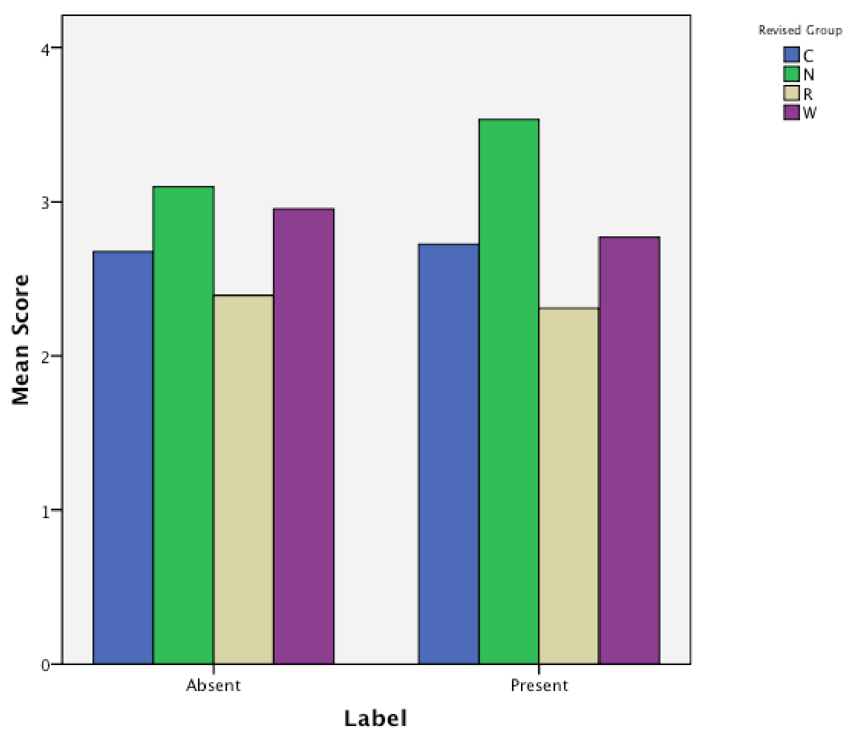


Figure 10. Similarity ratings by participant type and label condition

Note: Participant types are connoisseurs (C), novices (N), retailers (R), and winemakers (W).

Comments made by some participants as they completed the task suggested a possible explanation for the absence of a label effect. In some situations, it seemed that instead of a label increasing the perceived similarity of a pair of wines by activating a latent category, it appeared to shift the context of the similarity judgment leading to a more stringent standard for what was considered similar. Thus, for example, given two Sauvignon Blanc wines (and no label), the judgment of similarity tended to be fairly high, perhaps because the importance of the varietal to overall similarity was so apparent. When the label "Sauvignon Blanc Wines" was present, however, this limited the considered range. Instead of responding, "Ah, yes! They are both sauvignon blanc" (a response that would presumably increase perceived similarity), the response was more along the lines of, "Well, for sauvignon blanc wines, these two are really quite different."

This insight suggested several types of re-analysis of the data. For sauvignon blanc wines, one might not expect novices to have the same reaction as experts; even though the grape varietal is visible on the label, they would probably not have the frame of reference to compare two different instances within the category (though they might take into account regional information, also available on the label). Yet, when impact of label was evaluated for different categories of wine pairs (e.g., Sauvignon blanc vs. Red vs. High end and so on), there was little consistency in the response of the experts and novices. (The following pattern descriptions reflect general trends, not statistical significance.) Sometimes the two groups responded similarly with label increasing similarity judgments (high-end, sweet), decreasing them (California chardonnay, red), or having relatively little impact (California, party). Sometimes the groups responded differently.

Many other analyses (not reported here) were conducted to explore possibly confounding variables: effects of task order, individual categories, different item sets (A and B), and individual pairs of items. I examined each participant group individually; I collapsed all experts. I sought individual differences in response to label. The results were extremely variable and I could not discover any underlying pattern.

Discussion

The results of the similarity task are difficult to interpret. The anticipated increase in similarity judgments in response to the presence of a priming label did not materialize in any regular way. I am left with little explanation for these results, except to say that there was considerable variability. The source of the variability is puzzling, but not entirely intractable. Similarity, while an intuitive concept, has a long history as an elusive one. In addition to the challenge of determining *which kind* of similarity matters for a given task or judgment (Heit & Rubinstein, 1994; Medin et al., 1993), there are many contextual factors that also influence judgments (Medin & Goldstone, 1995). Medin and Goldstone state, "Our thesis is that similarity involves multiple predicates and that similarity statements of the form 'A is similar to B' are really shorthand for 'A is similar to B in respects C according to comparison process D, relative to some standard E mapped onto judgments by some function F for some purpose G.'" (1995). In the context of this judgment task, not only are the novices and different types of experts bringing different background knowledge and experience to the task, but they may also be extracting different information from the stimuli and inferring different constraints and standards from the pairs of wines and labels presented.

Category Membership and Typicality

The goal of the category membership task was to establish whether the varied categories generated in the sorting task were "true superordinates" (Ross & Murphy, 1999, p. 505) of the wines. The main question was whether wines rated as belonging to a variety of different *kinds* of categories. The alternative—that only taxonomic categories were true superordinates—would be disproved if wines received high ratings on script-based categories. This task also provided an opportunity to test whether wines were cross-classified. That is, would participants rate a single wine to be a good member of more than one category?

Methods

Design

The design was a mixed within- and between-subjects design, with category serving as the within-subjects factor and expert type as the between-subjects factor. Participants rated the category membership of the 40 wines by sorting the wine label cards into nine piles. The first eight piles indicated scores ranging from 0 to 7 (0 = Not a Member; 7 = Very Typical Member). The ninth pile was for wines they didn't feel they knew well enough to judge. Participants repeated the process up to 13 times, each time for a different category.

Materials

I laid out five sheets of dark green paper in front of each participant, all in landscape orientation. The "X" category was on its own sheet, but the four other sheets each had two headings. It was usually possible to array the sheets so the numbers ran from 0 to 7 from the participant's left to right. Four headings had descriptions below them; these were "NOT a member" (0), "Fairly Typical Member" (4), "Very Typical Member" (7), and "Don't Know Well

Enough to Judge" (X). Participants sorted the 40 laminated wine label cards from Session 1 into the nine different piles. In addition to the wine label cards, I also prepared three demo items with images of (1) a dump truck, (2) hot air balloons, and (3) a flower vase. The names of the thirteen categories were printed and pasted onto small cards.

Procedure

Before beginning the task, participants heard these instructions, which I adapted from Ross & Murphy (1999):

> I want you to evaluate each wine in terms of how good an instance of a category it is. Essentially, you will be giving each wine a rating from 0 to 7, where 0 indicates that the wine is not a member of the category, 3 indicates that it is a fairly good member of the category and 7 indicates that it is an excellent member of the category. Use the other numbers as needed. If you don't think you know a wine well enough to judge, put it here. [pointed to "X" mat]

I then brought out the demo cards to illustrate how to use the sorting mats.

> For example, if the category were <u>vehicles</u> and you had to judge how good an example of the category these items were [showed demo cards], most people would give the vase a 0 [while placing vase card on 0 mat], because it is not a vehicle. For this card, [showed dump truck card while placing] they might give a 7 or perhaps a 6. While the hot air balloon is technically a vehicle, it's not a very typical one, so it would probably get a lower score like a 1 or a 2. Note that some people might disagree, claiming that both the dump truck and the hot air balloon are both vehicles and equally good examples. There are no right or wrong answers here; I am interested in what you think.

I alerted participants to the fact that we would be evaluating a number of categories and encouraged them to move through the individual items fairly quickly.[35] However, four individuals (one connoisseur and three retailers) were unable to complete all twelve categories because of time constraints. One retailer skipped a single category (sweet); the other three individuals skipped three categories (champagne, high end, and sweet). No novices completed this task.

Categories

All participants rated the same category first, evaluating all of the wine labels for the membership in the category, "Wines." The other categories were shuffled within blocks.[36] The full list of categories (by block) is presented in Appendix R.

Results

There were two substantive objectives for the category membership task. The first was to verify that experts actually believed the categories to be descriptive of the wines. In other words, did experts rate at least some wines to be members of each of the category types? The extent to which experts agreed on what was and what was not a member of a category could provide support for the category's legitimacy. The second goal was to test whether wines can be thought

[35] I also asked them to say the number of the card and its score aloud so I could record their responses (both manually and on tape). When categories included a large number of similar scores (e.g., most of the wines were not category members), participants often chose to lump them as the default (e.g., "the rest are zeroes") and only say aloud the numbers and scores of category members.

[36] Because I could not predict the amount of time individuals needed to complete each sort, I could not expect everyone to complete all 13 categories. I did not want to use the same order across all subjects, but neither did I want to use a purely randomized category order because important categories might have missing data. My solution was to create subsets of categories, order the sets to prioritize those containing the most important categories, and shuffle the order of the items within sets for each participant.

to be good members of multiple categories. Ross and Murphy (1999) found that most food items were judged to be a good member of only one taxonomic-type category, but that they were often considered good members of multiple script categories. It is my intuition that for wines, cross-classification in the form multiple category memberships will be common and that this will not be limited to script categories. Finally, data from this task provided an index of typicality for the wine-category pairs used in the similarity task as well as for the premise wines used in the inference task.

Membership in Individual Categories

Were the categories used in this task considered meaningful descriptors of at least some wines? I certainly expected the answer to be yes; all of the categories[37] were drawn from the set of labels that the experts had created in the sorting task. Thus, at least one expert had spontaneously created these categories to organize the wines. However, because categories generated in production tasks do not always stand up in rating tasks (see Tversky & Hemenway, 1984), it was important to confirm that the experts actually found these categories meaningful.

The highest possible score was 7 ("a very typical member"), so I tallied the number of participants who gave at least one 7 score for each category. For five of the twelve categories (Italian, High End, Single Varietal, Sweet, and Champagne) every participant gave at least one wine the highest score. Although there are some categories where one or two individuals did not give any wine the top score, this was relatively infrequent. For "wines," Old World wines, and sparkling wines, a single retailer considered no wine in the set to be a very typical example. For Bordeaux blend wines, light wines, and sauvignon blanc wines, two experts found no very

[37] The one exception was the "wines" category.

typical examples. Even for the weakest category (party wines), 20 of the 24 experts considered at least one wine to be a very typical example. There did not appear to be large or systematic differences in terms of the types of categories for which experts could find excellent examples. These results suggest that yes, these categories were meaningful to the experts; most of them found at least one strong example of each category.

Table 8. Mean membership agreement

Category type	Category name	Connoisseurs	Retailers	Winemakers	All
role	Party	0.65	0.42	0.39	0.42
price/quality	High End	0.61	0.51	0.61	0.47
combo	Bordeaux Blend	0.86	0.64	0.63	0.67
style	Sweet	0.48	0.76	0.71	0.67
process	Single Varietal	0.76	0.68	0.73	0.70
combo	Light White	0.82	0.72	0.72	0.73
region	Old World	0.80	0.73	0.79	0.77
grape	Sauvignon Blanc	0.95	0.85	0.78	0.84
region	Italian	0.97	0.72	0.94	0.86
type	Sparkling	0.93	0.88	0.92	0.91
poly	Champagne	0.95	0.88	0.94	0.91
	Wines	1.00	1.00	1.00	1.00

Note: The category names are ordered by the average agreement across all participants.

To what extent did experts agree about which wines were members of the categories? Following Barsalou (1983), all 0 ratings (not a member) were recoded as -1, and the remaining scores (1-7) were coded as +1 to indicate membership. The absolute value of the average of these scores for a wine-category pair indicates the degree to which participants agreed about the membership of that wine in that category. Appendix S presents this data in detail; Table 8 presents the 40 agreement scores, averaged by group for each category.

All experts agreed that the wines were wines, and there was also high agreement about which wines were sparkling and which were champagne. The most "controversial" categories were the one included role category (party wines) and the price/quality category (high end wines). Examination of the average typicality ratings (see Appendix T) suggests that the group differences in agreement are sometimes due to differences in the stringency with which group members applied category membership criteria. For example, connoisseurs took a more strict interpretation of Bordeaux Blend and Sauvignon Blend categories than the other two groups. Retailers were generally the most flexible about category boundaries.

Wine Cross-classification

Are wines considered good members of <u>multiple</u> categories? Experts were certainly willing to assign wines to multiple categories (see Figure 11). Only four wines were considered members of just a single category, whereas a few wines were considered members of half the categories tested. On average, wines were considered to be members of 3.1 of the categories in the set. Given that time constraints limited the number of categories evaluated (more than 600 categories were generated in the sorting task, but only 12 were included in this task) this is surely a conservative estimate.

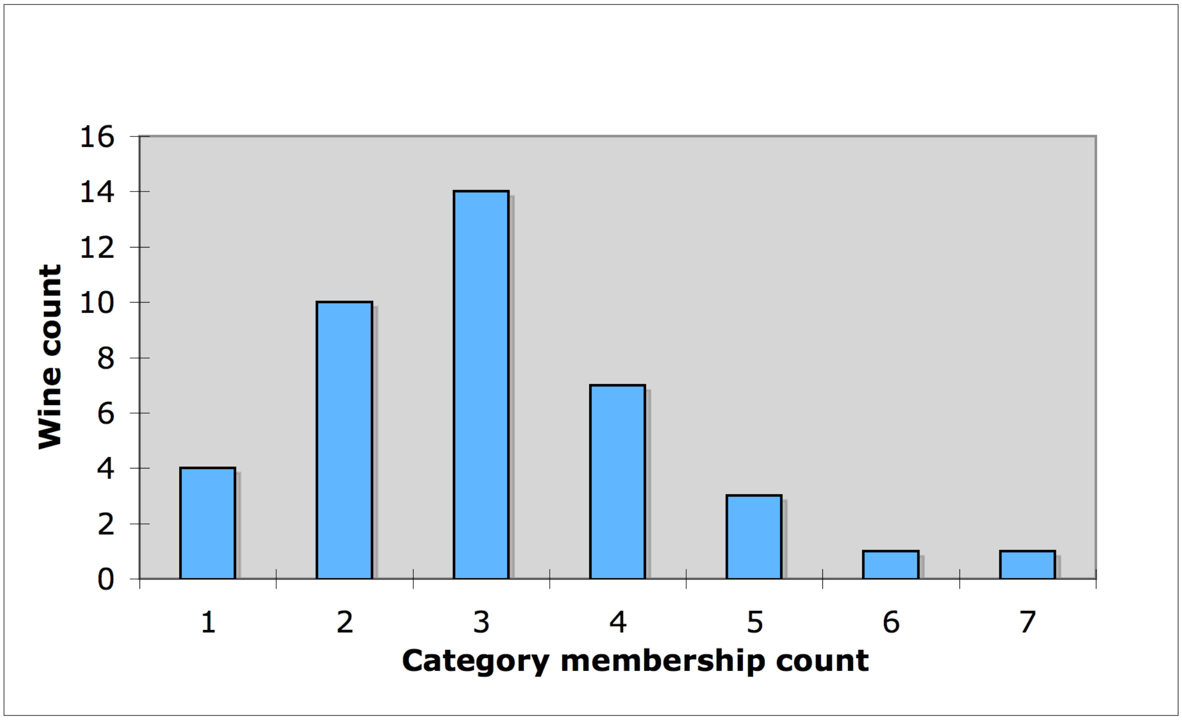


Figure 11. Category memberships of wines

Note: This histogram depicts how many wines were members of a given number of categories (ranging from one to seven). A wine was considered a member of a category if more than 75% of respondents identified it as a member of the category (i.e., gave it a non-zero rating).

Evaluation of Wine-Category Pairs from Similarity Task

A minor objective of the category membership task was to assess the appropriateness of the items used in the similarity task. Specifically, were the wines used for the different category sets (taxonomic, script, combination) in the similarity task roughly equivalent in terms of their typicality? I averaged the typicality scores for those wines (with respect to the appropriate category) by category type. In other words, for the wines in question, the typicality scores were averaged separately for the taxonomic (Sauvignon Blanc, Old World, and Italian), script (Party, Sweet, High End, Sparkling and Single-Varietal), and combination (Bordeaux Blend and Light White) sets. The results for the three category sets were within two points of each other, with the taxonomic items receiving somewhat higher ratings ($M = 6.36$) compared to the script ($M = 5.30$) and combination ($M = 4.96$) wine-category pairs. On average, the wines for each of the three sets were considered at least "fairly typical" of the categories on average.

Evaluation of Typicality of Premise Wines for Inference Task

The experts' ratings of how good an example each premise wine was of the category "Wines" served as an indicator of typicality. Examining only the four wines that would be used in the Inference task, a 3 x 4 ANOVA (Group by Premise Wine) showed both expert group ($F(2,83) = 10.894$, $p = .000$) and premise wine ($F(3, 83) = 3.367$, $p = .022$) to be significant, but there was no interaction. Retailers gave lower typicality ratings than the other two expert types, across all premise wines. All three groups rated the Taylor Lake Country Red (#14) as less typical than the other premise wines. Figure 12 shows the pattern of typicality scores.

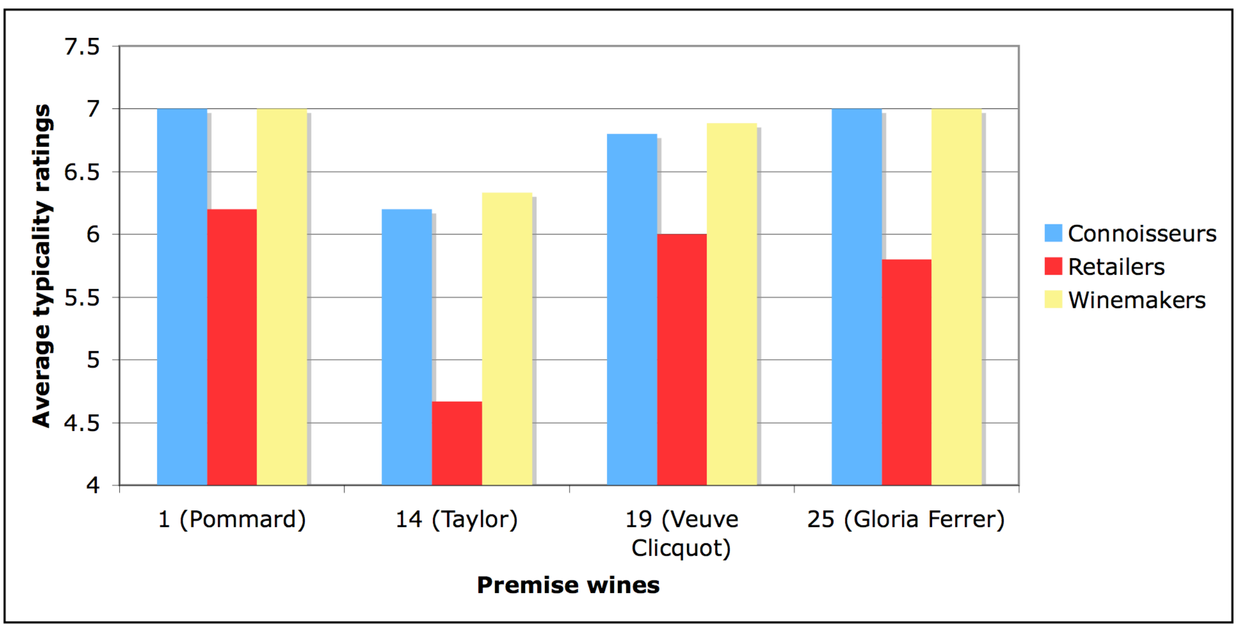


Figure 12. Mean typicality ratings for inference task premise wines

Note: The category membership task supplied the data for these ratings.

Discussion

This task established that wines are rated as belonging to more than just taxonomic categories. Most participants found excellent examples for all of the categories used in the task. However, the different types of categories were not equal. Agreement on category membership was more consistent for the taxonomic and combination categories included in the task. Script categories such as "party wines" were more often subject to differences of opinion. The rating task also firmly established that wines were cross-classified; the majority of wines in the sample were judged to be members of three or more of the examined categories. Thus the category membership task provided support for the legitimacy of the categories examined.

Inference

We now turn to the inference tasks, which assess the degree to which—and the way in which—experts use categories to make predictions about unfamiliar properties. One goal of this task was to determine whether the categories that experts generated in the sorting task were used to guide inference. The extent to which they do lends support for the psychological legitimacy of those types of categories. In addition, this task explored the influence of a number of factors on patterns of reasoning.

Sloman's feature based model of induction (1993) predicts that greater feature overlap between premise and conclusion categories leads to judgments of greater argument strength. Extrapolating to this task, it is reasonable to expect that experts will prefer to extend a novel property to target wines that share more features with the premise wine. There are theoretical reasons to expect that the *kind* of feature match matters. Research on inference has suggested that certain properties are more important for inference, specifically, those that play a more causal, or central role in the structure of the category. So, for example, deep features—those that other features depend upon—may play a greater role in classification decisions (Ahn, 1998, Sloman, Love & Ahn, 1998) and inference (Rehder, 2006).

In addition, there is reason to expect that the type of property may influence the types of categories that experts use on the inference task. Heit and Rubenstein (1994) showed that property type influenced the type of similarity being evaluated. For example, people were more likely to say that taxonomically related mammals shared anatomical properties, but expected behaviorally related animals to share behavioral properties. Similar phenomena have been observed for reasoning about food items (Ross & Murphy, 1999) and fish (Shafto & Coley, 2003). The present analyses will test the degree to which this also is true for wine experts.

Thus far, I've identified a number of factors that may affect reasoning: overall judgments of similarity based on feature matching, the kind of feature matching or category membership, and the type of property being extended. What about the impact of the premise? Research has shown that the similarity between a premise and the conclusion category has an influence on <u>whether</u> a property is extended to a target (e.g., Rips, 1975; Osherson et al., 1990), but does the premise itself have other kinds of influence as well? Does it influence the <u>kind</u> of similarity that is relevant? There are several reasons to think so. The relevance theory of induction argues that "people assume that the premises are informative with respect to the conclusions" (Medin et al., 2003, p. 517). Given two premise categories, people may seek a relationship between them and extrapolate from that. For example, given an argument in which cows and grass both possess a property, people may infer that the relevant relationship is a causal one: the grass has the property and the cows get it by eating the grass; therefore other animals that eat the same kind of grass as the cows do (or perhaps eat the cows themselves) may also have the property. When there is a single premise category, people may use what is distinctive about it to guide their induction.

Nelson and Miller (1995) observed a similar phenomenon for social categories. Information about a person was more likely to be generalized to others that shared a distinctive property with the person. Thus, preferences of a skydiver who was also a dog owner were thought more likely of other skydivers than of other dog owners. In other words, the distinctive properties of a person affected patterns of inference. Similarly, it may be true that what is distinctive about a wine is more likely to be retrieved and affect property extension.

Another reason to think that the premise will affect perception of what property matters comes from research on cross-classification. Murphy and Ross (1999) found that when an item could be classified in different ways, a variety of factors—even apparently superficial ones such as the format of information presentation—could affect which category guided inference. One potential explanation they forwarded was that when items can be cross-classified, people must first decide which category membership is relevant to the property.

And finally, assuming experts do use the categories from the sorting task for inference, do the three types of experts rely on them equally? Specifically, do the subtle differences in emphasis observed in the sorting task carry over to inference tasks?

Thus, some potentially important factors are individual, the property, the premise items, and the relations among them. The present analysis will explore these possibilities.

Methods

Design

The design was a mixed within- and between-subjects design, with property and premise serving as within-subjects factors and expert type as the between-subjects factor. Four of the 40 wines from Session 1 served as premises. All participants evaluated the same twenty items: five properties for each of the same four premise wines. Participants viewed all properties for one premise wine before proceeding to the next. The same presentation order (1-5, as listed below) was used for all premise wines and for all participants. The order of the wines varied. I created a list of all 24 possible permutations of the four wines, randomly assigned them to participants, and repeated the process to fill out the group.

Materials

Participants responded to the following five inference questions:

1. Suppose I were to tell you that this wine has Property Q. What other wines would you expect to have Property Q?
2. Imagine that a new type of chemical analysis has discovered that this wine has a chemical compound called "Xergia" in it. What other wines would you expect to have Xergia in them?
3. Imagine that a chef has prepared a novel dish. The flavor of the dish is enhanced by this wine. What other wines would you expect to go well with the dish?
4. Imagine that a new fungus called "Nerual" has been detected in this wine. What other wines would you want to test for "Nerual"?
5. Imagine that aliens have come to earth. They've tried some wines and they really adore this one. What other wines would you expect them to like?

Each of these texts was printed on white paper, cut out, and glued to the center of a letter-size piece of orange paper, landscape orientation. The same 40 wine label cards were used in all trials; one was pulled out to serve as the premise and the remainder were used as a checklist, as described below.

The four premise wines were (1) Louis Jadot Pommard (a relatively high end French pinot noir), (2) Taylor Lake Country Red (an inexpensive New York State hybrid blend), (3) Veuve Clicquot (a high end French champagne), and (4) Gloria Ferrer Chardonnay (a moderate to high end California Chardonnay).

Procedure

As instructions I told participants:

> For the next task, I'm going to describe a hypothetical situation and ask you to make a prediction, using limited information. There's no right or wrong answer; you should just make your best guess.

Next I showed them the first premise wine, turned over the first property prompt so they could follow along while I read the inference question aloud. The participants then looked through the set of 39 wine cards. After they finished making their selections, I asked them to explain their reasoning before proceeding to the next inference question. No novices completed this task.

Results

Data analyses focused on two main questions. First, were the categories that emerged in the initial sorts also used in the inference task? In particular, did common sorting categories get used for inference? Or did the demands of the sorting task yield an emphasis on categorical distinctions that have little value for making inferences?

Second, did different kinds of experts emphasize different category types in making inferences? In the sorting task, some mild differences emerged in experts' preferences for certain types of categories. Did these preferences have an impact on the reasoning strategies they subsequently used? Or did the task itself—the problem posed by a particular premise-property combination—prove to have more influence on experts' reasoning than any pre-existing preferences?

To answer these questions, I first examined the degree to which experts' sorts predicted the wines they expected to share novel properties. In other words, could distances from the sorting data predict choices on the inference task? I also explore the principles that might underlie the patterns of choices, examining both the similarity between the premise and target on key dimensions and the types of justifications participants offered. Finally, I present findings from some additional analyses that explore the impact of similarity and typicality on the main inference findings. Thus, analyses of two sources of information—experts' choices and their justifications of those choices—provide insight into experts' reasoning.

Choice data

Examination of which wines experts believed would share a property with the premise is a useful way to explore which categories were used, as well as whether experts differed in their approaches to the inference task and whether there were property effects. Rather than present an analysis of the total number of wines to which a property was projected (i.e., the overall number of property extensions), I examine the choice data in terms of the relationship between the targets (extensions) and the premises.[38]

Sorting to inference.

First, was there any evidence that the categories created in the sorts were influencing the choices made on the inference task? If the sort categories were related to the inference choices, then for a given premise wine, experts should be more likely to choose target wines that they had grouped with that premise during the sorting task. In other words, the lower the distance between the premise and the target, the more likely the target should be chosen. This logic predicts a negative correlation between the premise-target distance and the probability of a target wine being chosen for that premise wine.

---

[38] Analyses of total number of extensions are presented in Appendix U. I give them less emphasis here because pure counts of wine choices are heavily influenced by the particular items included in the set of wines and also are very sensitive to the impact of individuals who selected *all* wines.

Each expert made inferences about five different properties for each of four premise wines. I correlated whether the expert chose a target wine for a particular property-premise combination with the distance between that target and the premise wine, based on the expert's first sort. Thus, I conducted five different correlations for each individual, correlating the 156 distances separately with the 156 choices for each of the five properties. I aggregated the results and conducted one- sample t-tests to establish that the average correlations were negative. Table 9 presents the results.

Table 9. Mean correlation of premise-target distance (first sort) with target choice

	Mean	Std. Deviation	t	df
Aliens	-0.24^{**}	0.18	-6.57	(23)
Dish	-0.21^{**}	0.15	-6.88	(23)
Fungus	-0.23^{**}	0.27	-3.87	(21)
Q	-0.25^{**}	0.17	-7.12	(22)
X	-0.22^{**}	0.15	-6.89	(22)

$^{**}p < .01$, two-tailed

Averaged across all experts, there was a significant negative correlation between the likelihood of a target wine being selected as an extension of a property. This was true for all properties; a one-way within-subject ANOVA found no main effect for property. The closer a target wine was to a premise wine in an expert's first sort, the more likely that wine would be chosen to share any of the five types of properties on the inference problems. (This general pattern also held when these analyses were broken down by expert type. See Appendix V for details.)

Table 10. Mean correlation of premise-target distance (first sort) with target choice, aggregated by sort cluster

	A (Regional)			B (Evaluative)			C (Color-Grape Combo)		
	M	*SD*	*t (df)*	*M*	*SD*	*t (df)*	*M*	*SD*	*t (df)*
Aliens	**-0.26****	0.13	-6.20 (8)	**-0.15****	0.07	-5.29 (5)	**-0.28***	0.25	-3.32 (8)
Dish	**-0.19****	0.07	-7.84 (8)	**-0.17***	0.14	-2.88 (5)	**-0.26****	0.20	-3.80 (8)
Fungus	**-0.42****	0.20	-6.44 (8)	**-0.03**	0.14	-0.41 (5)	**-0.11**	0.27	-1.28 (8)
Q	**-0.29****	0.14	-6.29 (8)	**-0.08**	0.15	-1.37 (5)	**-0.32****	0.13	-6.86 (7)
X	**-0.22****	0.15	-4.32 (8)	**-0.09**	0.12	-1.81 (5)	**-0.31****	0.10	-8.56 (7)

$^{**}p < .01$, two-tailed
$^{*}p < .05$, two-tailed

Examination of these correlations by *type of sort* revealed a slightly different pattern. The A-B-C designations had emerged in the sorting analyses as more potent clusters than expert type. When the correlations were aggregated for these clusters, property did matter. A two-way mixed ANOVA (Property x Sort Cluster) found a significant effect for sort cluster, $F(2, 17) = 8.20$, $p = .00$, and a significant property by sort cluster interaction, $F(8, 68) = 2.69$, $p = .045$. See Table 10. Type B (evaluative) sorts were less predictive of inference choices than the other two sort types (post hoc pairwise comparisons yielded ps $< .01$, with Bonferroni adjustments). However, although they were weak, there were significant negative correlations between distances based on the type B (evaluative) sorts and inference choice for the Aliens and Dish properties. Differential impact on inferences about the fungus property also contributed to the property by sort cluster interaction. A one-way, between-subjects ANOVA for the fungus property alone found a significant effect of sort cluster, $F(2,19) = 6.37$, $p = 01$. Post-hoc pairwise comparisons showed that the mean correlation for type A (regional) sorts was stronger than for the other two sort types (*p*s < .05, with Bonferroni adjustments).

Thus, based solely on the sorting data and inference choices there is some evidence that the sorts are guiding inference; wines that were closer to a premise wine in an expert's first sort are more likely to be chosen. In addition, the types of groups created in certain sorts are better at predicting certain types of inferences.

Coding for Premise-Target Wine Matches

To examine the degree and type of similarity between the premise and target wines, I coded each wine in terms of whether it matched the premise wine on four dimensions: color, grape, price, and region. Except for price, these were the dominant category types generated in the sorting task. The PLCB database provided objective information about wines' values on these four dimensions, for almost all wines in the set.

For grape, price, and region, I calculated three levels of matches. For grape, wines matched if they: used only the same varietal (or blend of varietals) as the premise (L3), used all of the grapes (and possibly others) as the premise (L2), or used any of the grapes that were in the premise (L1). For region, wines matched if they were: from the same sub-country region (such as a state; L3), from the same country (L2), or from the same larger region (L1). For price, wines matched if they were: within \$5.00 of PLCB price (L3), within \$10.00 (L2), or within \$25.00 (L1). For color/type, there was only one level: to be considered a match, a wine had to be the same color (red/white) and the same type (sparkling or still). Table 11 shows the specific standards for each premise wine.

Table 11. Criteria used to calculate matches

Match type	Level	Louis Jadot Pommard (1)	Taylor Lake Country Red (14)[39]	Veuve Clicquot (19)	Gloria Ferrer Chardonnay (25)
Grape	1	Pinot noir (only)	-------------	Pinot noir, pinot meunier, <u>and</u> chardonnay (only)	Chardonnay (only)
	2	Pinot noir (at least)	-------------	Pinot noir, pinot meunier, <u>and</u> chardonnay (at least)	Chardonnay (at least)
	3	Same as L2	-------------	Pinot noir <u>or</u> pinot meunier <u>or</u> chardonnay (at least)	Same as L2
Region	1	Burgundy, France	New York, US	Champagne, France	California, US
	2	France	US	France	US
	3	Old World	New World	Old World	New World
Price	1	$24.99-$34.99	$0.00-$9.99	$38.99-$48.99	$11.99-$21.99
	2	$19.99-$39.99	$0.00-$14.99	$35.99-$55.99	$6.99-26.99
	3	$4.99-$54.99	$0.00-$29.99	$18.99-$68.99	$0.00-$41.99
Color/ Type	1	Red/Still	Red/Still	Sparkling	White/Still

This matching scheme provides a rough indication of the similarity of targets to premises along several key dimensions. The match scores were used in two types of analyses, described in separate sections below. In the first section, I examine the proportions of correct choices by match type, averaged across all target wines and levels of match. In the second section I examine

[39] The exact varietals used to make #14 are not printed on the label. Participant comments, however, suggested that the wine blends a variety of reds including vitis labrusca grapes and hybrids. These grapes were rarely present in other wines in the set, therefore no other wines were likely to use the same exact blend of grapes.

the proportions of specific target choices, supplemented with information about the degree and type of match for each premise.

Specifically, it is my expectation that different categories—as evidenced by the different match types—will be important for different kinds of inferences. In particular, I predict that the aliens and dish properties will be more likely to depend on categories that are closely related to taste. Although I was unable to measure match based on style or flavor principles, color and grape have been shown to be indicators of flavor (Brochet & Dubourdieu, 2001; Hughson, 2003; Solomon, 1997). For fungus and xergia, either biological or ecological factors could be important. Of the available match types, that could lead to greater match scores for grape or region. Generally, I expect property to play a greater role in the type of match that matters than expert type, given the weak influence of expert type on the kinds of categories emphasized during the sorting task. On the other hand, when the property is blank (e.g., Property Q), the predictions are less clear. This could be where expert type has the greatest influence, where premise plays the greatest role, or where overall similarity matters most.

Match proportions

The target wines that a participant chose for a given inference trial (each of 20 premise-property combinations) were compared to the potential matches to calculate the number of hits (matches chosen), misses (matches not chosen), false alarms (non-matches chosen) and correct rejections (non-matches not chosen). The proportion of the 39 possible selections that represented match-consistent responses (the sum of hits and correct rejections) served as an individual's score for each trial. This transformation reduced the impact of individuals who chose *all* wines (or *no* wines) as well as the problem of some premises that had no matches on a dimension.

The scores were analyzed with a mixed between- and within-subjects 3 x 4 x 5 x 4 ANOVA (Expert Group x Premise Wine x Property x Match type).[40] For this analysis, I was most interested in the interactions with match type: Did the type of match that mattered depend on the type of property or on the wine given as the premise? Did these interact with the type of expert making the judgment? Was there evidence that the experts relied more heavily on a particular type of match in their reasoning? See Tables 12 and 13 for the ANOVA results. There were significant main effects for match type and premise, as well as significant interactions of match type with property and match type with premise.

Table 12. Between-subjects effects for inference task analysis of variance

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	860.790	1	860.790	3664.707	.000
RevisedGroup	.028	2	.014	.060	.942
Error	4.933	21	.235		

[40] To reduce the complexity of this analysis I analyzed the match scores for only the intermediate levels (L2).

Table 13. Within-subjects effects for inference task analysis of variance (with Greenhouse-Geisser adjustments)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
premise	8.603	1.899	4.530	65.512	.000
premise * RevisedGroup	.205	3.799	.054	.779	.540
Error(premise)	2.758	39.885	.069		
Property	1.308	2.291	.571	2.273	.107
Property * RevisedGroup	1.744	4.582	.381	1.516	.207
Error(Property)	12.082	48.114	.251		
Matchtype	11.935	1.711	6.975	133.004	.000
Matchtype * RevisedGroup	.076	3.423	.022	.422	.763
Error(Matchtype)	1.884	35.937	.052		
premise * Property	.544	4.485	.121	1.460	.216
premise * Property * RevisedGroup	.863	8.970	.096	1.158	.331
Error(premise*Property)	7.824	94.183	.083		
premise * Matchtype	8.839	3.266	2.706	48.274	.000
premise * Matchtype * RevisedGroup	.258	6.532	.040	.705	.658
Error(premise*Matchtype)	3.845	68.586	.056		
Property * Matchtype	.621	3.345	.186	3.566	.015
Property * Matchtype * RevisedGroup	.736	6.690	.110	2.115	.056
Error(Property*Matchtype)	3.655	70.247	.052		
premise * Property * Matchtype	.291	6.244	.047	.987	.438
premise * Property * Matchtype * RevisedGroup	.960	12.488	.077	1.632	.087
Error(premise*Property*Matchtype)	6.180	131.119	.047		

Overall, the type of match mattered (Figure 13), with grape match scores the highest and price match scores the lowest. Region and color were lower than grape and higher than price. (These differences were all significant by post-hoc pair-wise comparisons, *p*s < .05 using Bonferroni adjustments.)

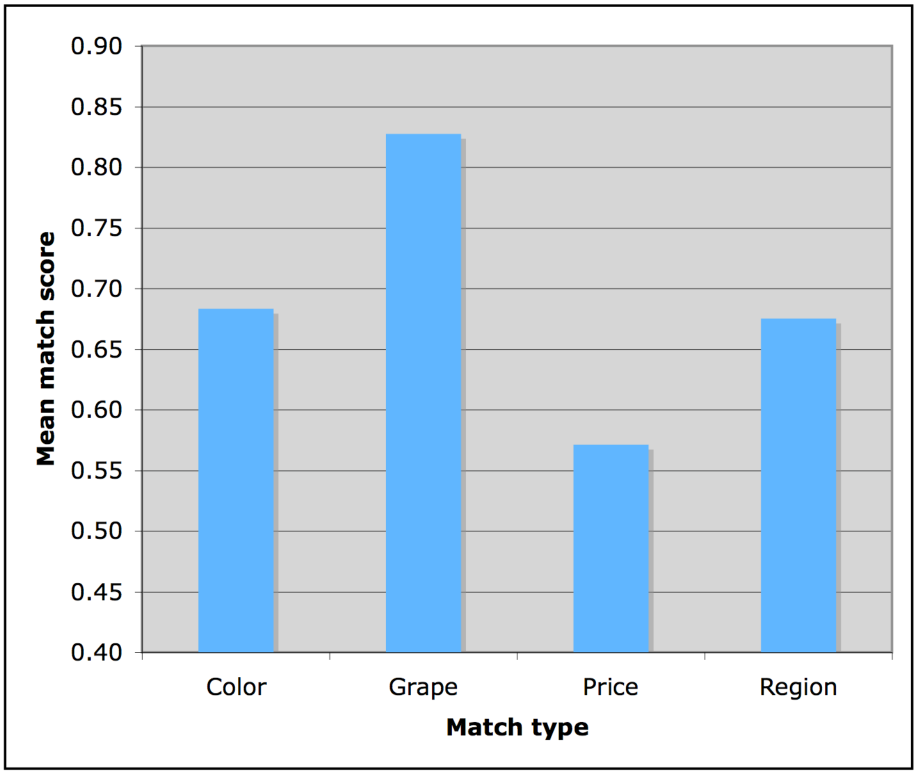


Figure 13. Mean match score by type of match

Note: This chart presents mean scores for each match type, averaged across all participants, properties, and premises. The scores are the proportion of selections (of the 39 choices for each trial) that represented a match-consistent (L2) response—either the selection of a wine that matches the premise or the rejection of a wine that does not match the premise.

Three of the four match types (grape, color, and price), showed comparable patterns of scores across the five properties: fairly consistent means for aliens, dish, Q, and xergia, with a lower mean for the fungus property. For the region-based matches, however, there was no difference across the five property types. For region matches, there was no dip for the fungus property; see Figure 14.

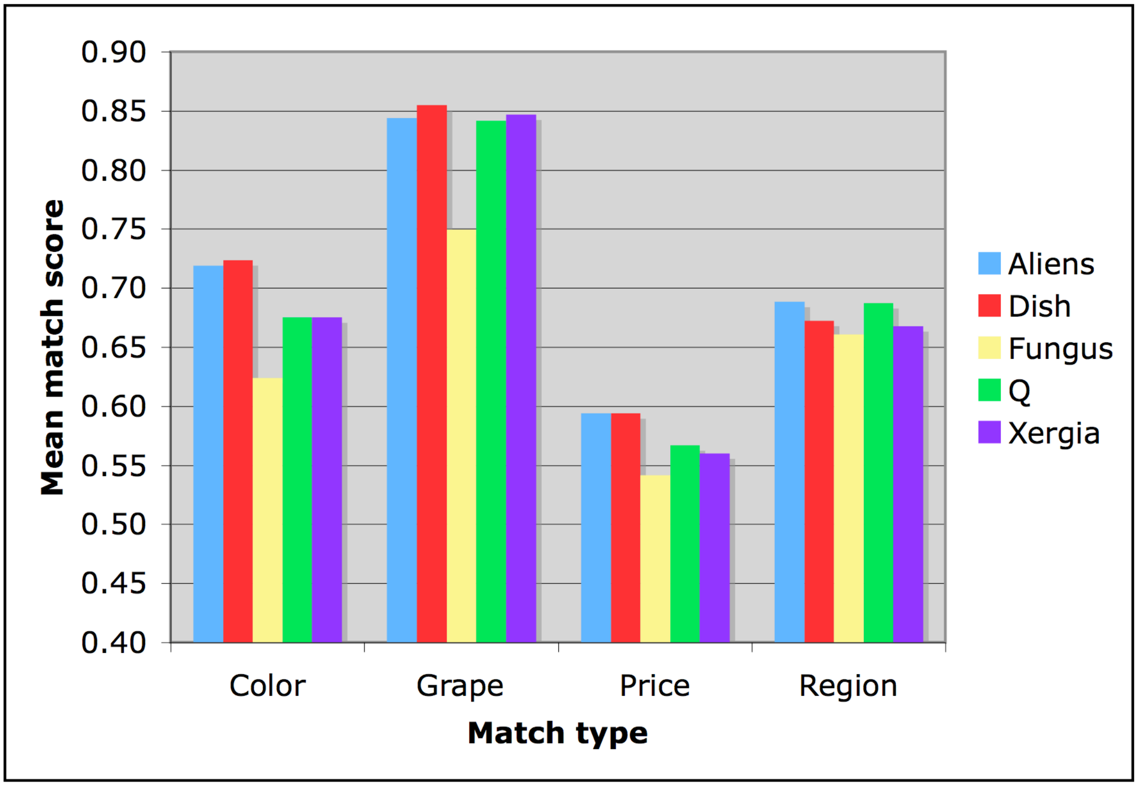


Figure 14. Mean match score: Property by match type

Note: This chart presents mean scores for each match type, in terms of the interaction between match type and property, averaged across all participants and premises. The scores are the proportion of selections (of the 39 choices for each trial) that represented a match-consistent (L2) response—either the selection of a wine that matches the premise or the rejection of a wine that does not match the premise.

The premise by match type interaction is presented in Figure 15. Grape scores were high for all premises, but the utility of the other match types depended on the premise. For the two French wines (1 & 19), price and region were good predictors of choices, but that was less true for the two U.S. wines (14 & 25). This national difference is probably because country (the region match at L2) is less meaningful for U.S. wines relative to French wines. See the discussion for more detail on this point. Color/type match was most important for the champagne (19) and the chardonnay (25), but less predictive for the two reds (1 & 14). [41]

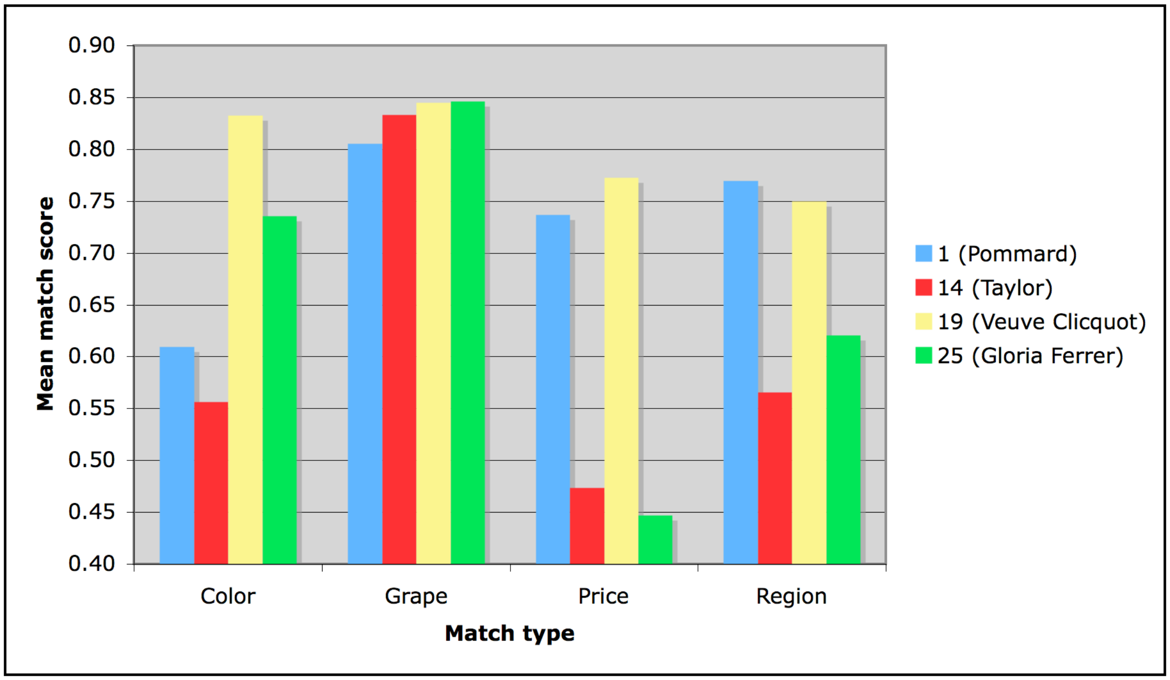


Figure 15. Mean match score: Premise by match type

Note: This chart presents mean scores for each match type, in terms of the interaction between match type and premise, averaged across all participants and properties. The scores are the proportion of selections (of the 39 choices for each trial) that represented a match-consistent (L2) response—either the selection of a wine that matches the premise or the rejection of a wine that does not match the premise.

Analyses of the inference choices thus far have shown that wines that were closer to the premise wine in a participant's first sort are more likely to be chosen to share properties with that premise. Although the correlations were not extremely large, they were significant and of moderate size. Examination of the role of property and premise on the type of category matches that were strongest showed that grape varietal was the most important type of match, across all

[41] Appendix W presents the main effect of premise.

properties and premises. There was some evidence of a property effect; for fungus, matches tended to be weaker, except for region matches. Finally, there was good evidence that the premise affected the type of matches that mattered, but this may have been an artifact of the coding system. No reliable effects involving expert type emerged.

Choices by target.

The inference results presented have used transformed and processed data to facilitate analysis. It can also be helpful to review the data in a format closer to participants' actual response patterns. These results are presented in Tables 14 through 17. Each table shows the choices for one of the four premise wines, summarized as the proportion of each expert group that extended a property to a target wine. The left side of the table presents choices for target wines 1-20, the right side presents target wines 21-40. The null results for the premise wine are shaded out. The results are organized in columns by property. For each target wine row, there are three sub-rows: connoisseur picks (C), retailer picks (R), and winemaker picks (W); the shaded center column identifies the group. On either side of this center column are codes indicating the degree to which the target and premise matched on four dimensions: color/type (C), grape varietal (V), price ($), and region (R). The closer the match, the more codes, so a pair coded $$$ was closer in price than one with a single $. Target wine names and characteristics appear in Appendix J.

Choices are presented as the proportion of group members that chose that wine, given that premise, for that property. So, for example, when told that the Louis Jadot Pommard (#1) had the chemical compound Xergia, 33% (3 of 9) winemakers thought the Rothschild (#5) would, too. The match codes "CRR" indicate that wine #1 and wine #5 are the same color/type (non-sparkling reds) and are from the same country (France). The absence of any $ or V codes indicates that they are not similar in price (the price difference exceeded $25) and do not use similar grapes (one is a pinot noir, the other is a cabernet sauvignon blend). These match summaries provide a relatively quick way of reviewing the extent to which a target wine and a premise wine are similar on these key dimensions. To facilitate review of the 600 values on each table, proportions appear in bold if at least two-thirds of a group chose the wine, and are grayed out if one third or fewer chose it; the rest appear in normal font.

Table 14. Inference choices for Louis Jadot Pommard (#1) premise trials

Aliens	Dish	Fungus	Q	X	#	Matches	Group	Matches	#	Aliens	Dish	Fungus	Q	X
							C			0.60	0.40	0.20	0.20	0.00
					1		R	C$	21	0.00	0.20	0.20	0.10	0.30
							W			0.44	0.33	0.44	0.44	0.33
0.00	0.00	0.20	0.00	0.00			C			**0.80**	**0.80**	**1.00**	**0.80**	**0.80**
0.00	0.10	0.00	0.10	0.20	2	$	R	CGGG$RR	22	0.50	0.60	0.50	0.40	**0.70**
0.00	0.00	0.22	0.11	0.00			W			**0.67**	0.56	**0.78**	**0.67**	**0.89**
0.00	0.00	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.00	0.10	0.10	0.00	0.20	3	$	R	GG$$$	23	0.00	0.10	0.00	0.10	0.30
0.00	0.00	0.33	0.11	0.11			W			0.00	0.00	0.33	0.11	0.00
0.20	0.20	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.00	0.00	0.10	0.00	0.30	4	$$$	R	$	24	0.10	0.10	0.00	0.00	0.20
0.00	0.00	0.11	0.11	0.00			W			0.00	0.00	0.33	0.11	0.00
0.40	0.00	0.60	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.40	0.20	0.40	0.20	0.30	5	CRR	R	$	25	0.20	0.10	0.10	0.10	0.20
0.44	0.22	0.56	0.44	0.33			W			0.00	0.00	0.33	0.11	0.11
0.00	0.00	0.20	0.00	0.00			C			0.60	0.60	0.20	0.00	0.00
0.00	0.10	0.00	0.10	0.20	6	$R	R	C$	26	0.20	0.30	0.20	0.20	0.30
0.00	0.00	0.22	0.11	0.00			W			0.22	0.11	0.33	0.33	0.22
0.20	0.00	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.20	0.20	0.30	0.10	0.30	7	C	R	$	27	0.00	0.00	0.00	0.00	0.20
0.22	0.11	0.33	0.22	0.22			W			0.00	0.00	0.22	0.11	0.00
0.20	0.20	0.20	0.00	0.00			C			0.60	**0.80**	0.20	0.00	0.00
0.60	0.60	0.30	0.20	0.40	8	C$$$R	R	C$	28	0.00	0.00	0.10	0.00	0.30
0.44	0.22	0.33	0.22	0.11			W			0.22	0.11	0.33	0.33	0.22
0.80	**0.80**	0.40	0.40	0.60			C			0.20	0.00	0.20	0.00	0.00
0.40	**0.80**	0.30	0.40	0.60	9	CGGG$	R	C$	29	0.00	0.00	0.10	0.00	0.30
0.44	0.44	**0.67**	**0.78**	**1.00**			W			0.22	0.00	0.33	0.22	0.22
0.00	0.00	0.20	0.00	0.00			C			0.40	0.20	0.60	0.20	0.00
0.00	0.10	0.20	0.00	0.20	10	$R	R	C$RR	30	0.10	0.00	0.20	0.00	0.30
0.00	0.00	0.22	0.11	0.00			W			0.44	0.00	0.56	0.22	0.22
0.00	0.00	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.00	0.00	0.00	0.00	0.20	11		R	$	31	0.00	0.00	0.10	0.00	0.30
0.00	0.00	0.22	0.11	0.00			W			0.00	0.00	0.22	0.11	0.00
0.00	0.00	0.20	0.00	0.00			C			0.40	0.20	0.20	0.00	0.00
0.00	0.10	0.00	0.00	0.20	12	$	R	C$$$R	32	0.30	0.20	0.40	0.20	0.30
0.00	0.00	0.22	0.11	0.00			W			0.33	0.22	0.33	0.33	0.11
0.00	0.00	0.60	0.20	0.20			C			0.00	0.00	0.40	0.00	0.00
0.40	0.20	0.60	0.40	0.30	13	$$RRR	R	$RR	33	0.30	0.10	0.30	0.20	0.20
0.22	0.00	0.44	0.22	0.00			W			0.11	0.00	0.44	0.22	0.00
0.00	0.00	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.00	0.00	0.00	0.00	0.20	14	C$	R	$	34	0.00	0.10	0.00	0.00	0.20
0.00	0.00	0.11	0.11	0.00			W			0.00	0.00	0.22	0.11	0.00
0.00	0.00	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.00	0.00	0.10	0.00	0.20	15	$R	R	C	35	0.00	0.00	0.20	0.00	0.20
0.00	0.00	0.33	0.11	0.00			W			0.22	0.22	0.33	0.33	0.33
0.20	0.20	0.20	0.00	0.00			C			0.40	0.60	0.20	0.00	0.00
0.20	0.30	0.40	0.20	0.30	16	C$R	R	C$	36	0.20	0.30	0.20	0.10	0.30
0.56	0.22	0.44	0.22	0.11			W			0.22	0.11	0.33	0.33	0.22
0.00	0.00	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.00	0.10	0.00	0.00	0.20	17	$	R	$R	37	0.00	0.00	0.00	0.00	0.20
0.00	0.00	0.22	0.11	0.00			W			0.00	0.00	0.33	0.11	0.00
0.40	0.20	0.20	0.00	0.00			C			0.60	**0.80**	0.20	0.00	0.00
0.20	0.30	0.20	0.20	0.30	18	C$R	R	C$	38	0.00	0.00	0.10	0.00	0.30
0.44	0.22	0.33	0.33	0.22			W			0.33	0.11	0.33	0.33	0.22
0.00	0.00	0.60	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.40	0.30	0.30	0.30	0.30	19	GG$RR	R		39	0.10	0.20	0.10	0.10	0.20
0.00	0.00	0.44	0.11	0.00			W			0.00	0.00	0.33	0.11	0.11
0.00	0.00	0.20	0.00	0.00			C			0.60	0.40	0.20	0.00	0.00
0.00	0.10	0.10	0.10	0.20	20	$	R	C$$	40	0.00	0.30	0.20	0.20	0.30
0.00	0.00	0.33	0.11	0.11			W			0.22	0.11	0.33	0.33	0.11

Table 15. Inference choices for Taylor Lake Country Red (#14) premise trials

Aliens	Dish	Fungus	Q	X	#	Matches	Group	Matches	#	Aliens	Dish	Fungus	Q	X
0.20	0.40	0.20	0.00	0.00			C			0.20	0.40	0.20	0.00	0.20
0.10	0.20	0.10	0.10	0.20	1	C$	R	C$R	21	0.10	0.30	0.10	0.00	0.20
0.00	0.00	0.33	0.22	0.00			W			0.00	0.00	0.33	0.33	0.11
0.00	0.00	0.20	0.00	0.00			C			0.40	0.60	0.20	0.20	0.20
0.10	0.10	0.10	0.00	0.10	2	$R	R	C$$$	22	0.30	0.40	0.00	0.00	0.20
0.00	0.00	0.33	0.22	0.00			W			0.00	0.11	0.33	0.33	0.00
0.00	0.00	0.20	0.00	0.00			C			0.00	0.00	0.40	0.20	0.20
0.10	0.00	0.10	0.00	0.10	3	$$RR	R	$RRR	23	0.10	0.10	0.50	0.50	0.30
0.00	0.00	0.33	0.22	0.00			W			0.22	0.22	0.56	0.22	0.00
0.20	0.00	0.40	0.00	0.00			C			0.20	0.00	0.20	0.00	0.00
0.10	0.00	0.40	0.50	0.60	4	RR	R	$$RR	24	0.20	0.10	0.00	0.10	0.20
0.33	0.11	**0.67**	0.33	0.11			W			0.22	0.11	0.33	0.22	0.00
0.20	0.40	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.30	0.10	0.20	0.10	0.20	5	C	R	$RR	25	0.10	0.00	0.00	0.10	0.10
0.00	0.00	0.33	0.33	0.11			W			0.00	0.00	0.33	0.22	0.00
0.20	0.00	0.20	0.00	0.00			C			0.20	0.20	0.20	0.00	0.00
0.10	0.10	0.00	0.00	0.20	6	$$	R	C$R	26	0.40	0.30	0.10	0.00	0.20
0.11	0.22	0.33	0.22	0.00			W			0.00	0.00	0.33	0.22	0.00
0.40	0.60	0.40	0.20	0.20			C			0.20	0.00	0.40	0.20	0.20
0.10	0.20	0.20	0.10	0.20	7	CRR	R	$$RR	27	0.10	0.10	0.50	0.40	0.50
0.00	0.00	0.33	0.44	0.11			W			0.22	0.22	0.56	0.22	0.00
0.20	0.40	0.20	0.00	0.00			C			0.40	0.60	0.40	0.20	0.20
0.10	0.20	0.10	0.10	0.20	8	C$	R	C$$$RR	28	0.50	**0.80**	0.10	0.20	0.20
0.00	0.00	0.33	0.33	0.11			W			0.00	0.00	0.33	0.33	0.11
0.40	0.60	0.40	0.20	0.00			C			**0.80**	**1.00**	0.40	0.20	0.20
0.40	0.40	0.00	0.00	0.20	9	C$$RR	R	C$$RR	29	0.40	0.50	0.20	0.20	0.20
0.00	0.11	0.33	0.22	0.00			W			0.00	0.00	0.33	0.44	0.11
0.00	0.20	0.20	0.00	0.00			C			0.20	0.40	0.20	0.00	0.00
0.10	0.00	0.00	0.00	0.10	10	$$	R	C$$	30	0.20	0.40	0.10	0.00	0.20
0.11	0.11	0.33	0.22	0.00			W			0.00	0.00	0.33	0.44	0.11
0.40	0.00	0.20	0.00	0.00			C			0.20	0.00	0.20	0.00	0.00
0.30	0.10	0.10	0.30	0.30	11	RR	R	$$R	31	0.10	0.00	0.10	0.10	0.20
0.44	0.22	0.33	0.33	0.00			W			0.33	0.11	0.44	0.33	0.00
0.40	0.00	0.40	0.00	0.00			C			0.20	0.20	0.20	0.00	0.00
0.20	0.10	0.40	0.60	0.60	12	$$$RR	R	C	32	0.20	0.20	0.20	0.10	0.20
0.44	0.33	0.56	0.22	0.11			W			0.00	0.00	0.33	0.33	0.11
0.00	0.00	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.10	0.10	0.10	0.20	0.10	13	$	R	$$	33	0.10	0.10	0.00	0.10	0.10
0.00	0.00	0.33	0.22	0.00			W			0.00	0.00	0.33	0.22	0.00
							C			0.00	0.00	0.20	0.00	0.00
					14		R	$$RR	34	0.10	0.00	0.00	0.00	0.10
							W			0.00	0.00	0.33	0.33	0.00
0.00	0.00	0.20	0.00	0.00			C			0.60	0.60	0.60	0.60	0.40
0.10	0.20	0.00	0.10	0.20	15	$$	R	CRR	35	0.30	0.40	0.50	0.50	0.50
0.11	0.00	0.33	0.22	0.00			W			0.00	0.00	0.56	0.33	0.11
0.40	0.60	0.20	0.20	0.00			C			0.20	0.40	0.20	0.00	0.00
0.20	0.40	0.10	0.00	0.20	16	C$$$	R	C$R	36	0.40	0.30	0.10	0.00	0.20
0.00	0.00	0.33	0.33	0.11			W			0.00	0.11	0.33	0.33	0.11
0.00	0.00	0.20	0.00	0.00			C			0.20	0.00	0.20	0.00	0.00
0.10	0.10	0.00	0.10	0.10	17	$$$RR	R	$$$	37	0.10	0.20	0.00	0.10	0.20
0.00	0.00	0.33	0.33	0.00			W			0.22	0.22	0.33	0.22	0.00
0.20	0.40	0.20	0.00	0.00			C			0.20	0.40	0.40	0.00	0.20
0.20	0.40	0.00	0.00	0.10	18	C$$	R	C$$$R	38	0.50	**0.70**	0.10	0.00	0.20
0.11	0.11	0.33	0.33	0.11			W			0.00	0.00	0.33	0.44	0.11
0.00	0.00	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.10	0.20	0.10	0.20	0.10	19		R	RR	39	0.10	0.00	0.00	0.00	0.10
0.00	0.00	0.33	0.22	0.00			W			0.00	0.00	0.33	0.22	0.00
0.00	0.00	0.20	0.00	0.00			C			0.20	0.40	0.20	0.00	0.00
0.10	0.00	0.00	0.00	0.10	20	$$RR	R	CR	40	0.10	0.20	0.10	0.00	0.20
0.00	0.00	0.33	0.22	0.00			W			0.11	0.11	0.33	0.33	0.11

Table 16. Inference choices for Veuve Clicquot Champagne (#19) premise trials

Aliens	Dish	Fungus	Q	X	#	Matches	Group	Matches	#	Aliens	Dish	Fungus	Q	X
0.20	0.00	0.60	0.60	0.40			C			0.00	0.00	0.20	0.00	0.00
0.20	0.30	0.40	0.30	0.30	1	V$RR	R		21	0.00	0.00	0.00	0.10	0.20
0.00	0.00	0.44	0.22	0.00			W			0.00	0.00	0.33	0.11	0.00
0.40	0.20	0.40	0.00	0.00			C			0.00	0.00	0.40	0.40	0.40
0.10	0.20	0.00	0.20	0.20	2		R	VRR	22	0.00	0.00	0.20	0.00	0.30
0.00	0.11	0.33	0.11	0.11			W			0.00	0.00	0.44	0.22	0.33
0.00	0.40	0.60	0.40	0.60			C			**0.80**	**0.80**	0.60	0.40	0.60
0.00	0.30	0.00	0.00	0.30	3	V	R	CVVV$	23	0.40	0.60	0.20	0.60	**0.80**
0.00	0.11	0.44	0.22	0.44			W			**1.00**	**1.00**	**0.89**	**1.00**	**0.89**
0.00	0.00	0.20	0.00	0.00			C			0.00	0.00	0.20	0.20	0.20
0.10	0.10	0.10	0.10	0.30	4	$$	R		24	0.20	0.10	0.00	0.10	0.30
0.00	0.00	0.22	0.11	0.00			W			0.00	0.00	0.33	0.22	0.11
0.20	0.00	0.60	0.20	0.00			C			0.20	0.60	0.60	0.40	0.60
0.10	0.00	0.30	0.20	0.20	5	RR	R	V	25	0.10	0.20	0.00	0.20	0.30
0.00	0.00	0.33	0.11	0.00			W			0.00	0.00	0.44	0.22	0.56
0.00	0.20	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.00	0.20	0.30	0.10	0.30	6	R	R		26	0.10	0.00	0.00	0.20	0.20
0.00	0.00	0.33	0.11	0.00			W			0.00	0.00	0.33	0.11	0.00
0.00	0.00	0.20	0.00	0.00			C			0.00	0.20	0.20	0.00	0.00
0.20	0.10	0.00	0.30	0.20	7	$	R		27	0.10	0.00	0.00	0.10	0.20
0.00	0.00	0.33	0.11	0.00			W			0.00	0.00	0.33	0.11	0.00
0.00	0.00	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.10	0.20	0.10	0.20	0.20	8	$R	R		28	0.00	0.00	0.00	0.00	0.20
0.00	0.00	0.33	0.11	0.00			W			0.00	0.00	0.33	0.11	0.00
0.00	0.00	0.20	0.20	0.20			C			0.00	0.20	0.20	0.00	0.00
0.00	0.10	0.00	0.00	0.30	9	V	R		29	0.00	0.00	0.00	0.00	0.20
0.00	0.00	0.44	0.22	0.33			W			0.00	0.00	0.33	0.11	0.00
0.00	0.40	0.40	0.00	0.00			C			0.20	0.00	0.60	0.20	0.00
0.30	0.30	0.20	0.10	0.20	10	R	R	RR	30	0.00	0.00	0.30	0.00	0.20
0.00	0.11	0.33	0.11	0.00			W			0.00	0.00	0.33	0.11	0.00
0.00	0.00	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.10	0.00	0.10	0.00	0.30	11		R		31	0.10	0.20	0.10	0.10	0.30
0.00	0.00	0.33	0.11	0.00			W			0.00	0.00	0.33	0.11	0.00
0.00	0.00	0.20	0.00	0.00			C			0.20	0.00	0.20	0.20	0.00
0.00	0.10	0.00	0.00	0.20	12		R	$$R	32	0.20	0.10	0.10	0.30	0.20
0.00	0.00	0.33	0.11	0.00			W			0.00	0.00	0.33	0.11	0.00
0.40	0.60	**0.80**	**0.80**	**0.80**			C			0.20	0.40	**0.80**	0.40	0.40
0.20	0.40	0.40	0.30	0.30	13	V$RR	R	RR	33	0.30	0.40	0.40	0.40	0.20
0.11	0.22	0.56	0.11	0.44			W			0.11	0.11	0.44	0.11	0.11
0.20	0.20	0.20	0.00	0.00			C			0.40	0.20	0.40	0.00	0.00
0.10	0.00	0.00	0.00	0.20	14		R		34	0.10	0.20	0.00	0.10	0.20
0.00	0.00	0.22	0.11	0.00			W			0.00	0.00	0.33	0.11	0.00
0.40	0.60	0.40	0.20	0.40			C			0.00	0.00	0.20	0.00	0.00
0.40	0.30	0.30	0.60	0.50	15	CR	R		35	0.00	0.00	0.00	0.10	0.20
0.56	0.00	0.56	0.56	0.33			W			0.00	0.00	0.33	0.11	0.00
0.00	0.00	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.00	0.00	0.10	0.00	0.20	16	R	R		36	0.10	0.00	0.00	0.20	0.20
0.00	0.00	0.33	0.11	0.00			W			0.00	0.00	0.33	0.11	0.00
0.20	0.40	0.60	0.40	0.60			C			0.40	0.40	0.60	0.20	0.40
0.00	0.10	0.00	0.00	0.30	17	V	R	CR	37	0.20	0.20	0.20	0.50	0.50
0.00	0.11	0.44	0.22	0.44			W			0.33	0.00	0.56	0.56	0.33
0.00	0.00	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.00	0.00	0.10	0.00	0.20	18	R	R		38	0.00	0.00	0.00	0.00	0.20
0.00	0.00	0.33	0.11	0.00			W			0.00	0.00	0.33	0.11	0.00
							C			0.40	0.40	0.40	0.60	0.60
					19		R	V$	39	0.20	0.50	0.00	0.30	0.30
							W			0.00	0.00	0.44	0.22	0.56
0.20	0.40	0.60	0.40	0.60			C			0.00	0.00	0.20	0.00	0.00
0.00	0.30	0.00	0.20	0.30	20	V	R	$$$	40	0.00	0.00	0.00	0.20	0.20
0.00	0.11	0.44	0.22	0.56			W			0.00	0.00	0.33	0.11	0.00

Table 17. Inference choices for Gloria Ferrer Reserve Chardonnay (#25) premise trials

Aliens	Dish	Fungus	Q	X	#	Matches	Group	Matches	#	Aliens	Dish	Fungus	Q	X
0.00	0.00	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.30	0.20	0.10	0.00	0.10	1	$	R	$$$R	21	0.20	0.10	0.10	0.00	0.10
0.00	0.00	0.22	0.11	0.22			W			0.00	0.00	0.22	0.11	0.11
0.40	0.20	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.20	0.60	0.10	0.00	0.20	2	C$$$R	R	$	22	0.10	0.10	0.00	0.00	0.10
0.33	0.22	0.22	0.11	0.00			W			0.00	0.00	0.33	0.11	0.11
1.00	**0.80**	**1.00**	**0.80**	**1.00**			C			0.00	0.20	0.20	0.40	0.20
0.70	**0.70**	**0.70**	**0.80**	**1.00**	3	CVVV$$$RRR	R	VV$$RR	23	0.20	0.30	0.10	0.00	0.30
0.78	**0.89**	**0.78**	**1.00**	**1.00**			W			0.22	0.33	0.56	0.44	0.56
0.00	0.00	0.20	0.00	0.00			C			0.40	0.00	**0.80**	0.40	0.60
0.10	0.10	0.00	0.10	0.10	4	$RR	R	C$$$RRR	24	0.30	0.30	0.20	0.10	0.30
0.00	0.00	0.11	0.11	0.00			W			0.11	0.00	0.33	0.22	0.33
0.00	0.00	0.20	0.00	0.00			C							
0.30	0.10	0.10	0.10	0.10	5		R	**25**						
0.00	0.00	0.22	0.11	0.11			W							
0.20	0.20	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.20	0.20	0.00	0.00	0.20	6	C$$$	R	$$$R	26	0.30	0.10	0.10	0.00	0.10
0.11	0.00	0.22	0.11	0.00			W			0.00	0.00	0.22	0.11	0.11
0.00	0.00	0.40	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.30	0.10	0.40	0.10	0.10	7	RRR	R	C$$$RR	27	0.10	0.10	0.00	0.00	0.20
0.00	0.00	0.33	0.11	0.11			W			0.00	0.00	0.22	0.11	0.00
0.00	0.00	0.20	0.00	0.00			C			0.00	0.00	0.40	0.00	0.00
0.30	0.20	0.10	0.00	0.10	8	$$	R	$RRR	28	0.10	0.10	0.30	0.10	0.10
0.00	0.00	0.22	0.11	0.11			W			0.00	0.00	0.33	0.11	0.00
0.00	0.00	0.40	0.00	0.00			C			0.00	0.00	0.40	0.00	0.00
0.10	0.10	0.40	0.20	0.10	9	$$$RRR	R	$$RRR	29	0.10	0.10	0.30	0.00	0.10
0.00	0.00	0.22	0.11	0.11			W			0.00	0.00	0.33	0.11	0.00
0.20	0.20	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.40	0.30	0.00	0.00	0.20	10	C$$$	R	$$	30	0.10	0.10	0.00	0.00	0.10
0.11	0.11	0.22	0.11	0.00			W			0.00	0.00	0.22	0.11	0.11
0.00	0.00	0.40	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.00	0.10	0.30	0.00	0.10	11	RRR	R	$$$R	31	0.30	0.10	0.00	0.10	0.10
0.00	0.00	0.33	0.11	0.00			W			0.00	0.00	0.22	0.11	0.00
0.00	0.00	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.20	0.10	0.00	0.00	0.10	12	C$$RR	R	$	32	0.30	0.10	0.10	0.00	0.10
0.11	0.00	0.22	0.11	0.00			W			0.00	0.00	0.22	0.11	0.11
0.40	0.60	0.20	0.20	0.20			C			0.40	0.60	0.40	0.20	0.20
0.70	**0.90**	0.30	0.30	0.60	13	CVVV$$$	R	C$$$	33	0.40	**0.70**	0.10	0.00	0.20
0.56	**0.67**	**0.67**	0.44	0.56			W			0.44	0.44	0.33	0.22	0.22
0.00	0.00	0.20	0.00	0.00			C			0.40	0.20	0.20	0.00	0.00
0.10	0.10	0.00	0.00	0.10	14	$RR	R	C$$RR	34	0.30	0.60	0.10	0.00	0.20
0.00	0.00	0.11	0.11	0.00			W			0.44	0.33	0.33	0.22	0.11
0.00	0.00	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.10	0.10	0.00	0.00	0.10	15	$$$	R	RR	35	0.10	0.10	0.10	0.00	0.10
0.00	0.00	0.22	0.11	0.00			W			0.00	0.00	0.22	0.11	0.11
0.00	0.00	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.10	0.10	0.10	0.00	0.10	16	$$	R	$$$R	36	0.30	0.10	0.10	0.00	0.10
0.00	0.00	0.22	0.11	0.11			W			0.00	0.00	0.22	0.11	0.11
0.80	0.60	**1.00**	**0.80**	**0.80**			C			0.20	0.20	0.20	0.00	0.00
0.60	0.50	0.60	0.60	**0.80**	17	CVVV$$RRR	R	$	37	0.10	0.10	0.00	0.00	0.10
0.56	**0.67**	**0.78**	**0.78**	**0.89**			W			0.00	0.00	0.22	0.11	0.00
0.00	0.00	0.20	0.00	0.00			C			0.00	0.00	0.20	0.00	0.00
0.10	0.10	0.10	0.00	0.10	18	$$$	R	$$R	38	0.10	0.10	0.00	0.00	0.10
0.00	0.00	0.22	0.11	0.11			W			0.00	0.00	0.22	0.11	0.11
0.20	0.20	0.20	0.40	0.20			C			**1.00**	**1.00**	**0.80**	**0.80**	**0.80**
0.30	0.40	0.10	0.10	0.30	19	VV	R	CVVVRRR	39	**0.90**	**0.90**	**0.80**	**1.00**	**0.90**
0.22	0.33	0.56	0.33	0.44			W			**1.00**	**1.00**	**0.78**	**1.00**	**1.00**
1.00	**0.80**	**1.00**	**1.00**	**1.00**			C			0.00	0.00	0.20	0.00	0.00
0.90	**0.90**	**0.90**	**0.90**	**1.00**	20	CVVV$$RRR	R	$R	40	0.20	0.10	0.10	0.00	0.10
1.00	**1.00**	**1.00**	**1.00**	**1.00**			W			0.00	0.00	0.22	0.11	0.11

The Louis Jadot Pommard (#1, Table 14) is a French red burgundy made from pinot noir grapes. For this premise wine, strong grape varietal matches had the most impact on choice frequency. The most frequently chosen targets across all groups and properties are the two other pinot noir wines, with the French pinot noir (22) chosen somewhat more often than the California pinot noir (9), especially by connoisseurs. The two “methode champenois” wines (19 and 23, level two grape matches) were chosen moderately often, with more extensions to the French wine (19) than the New York state wine (23). Other evidence of mild regional influence were the moderate scores for fungus for a French red (5) and a French white from the same negociant as the premise (13). Somewhat surprisingly, for the dish property connoisseurs often chose two wines that do not appear to share much with the premise (28 and 38); these are the only merlots in the set. Justifications indicated that for connoisseurs merlots and pinot noirs share some meaningful stylistic similarities not captured by the matching criteria.

The Taylor Lake Country Red (#14, Table 15) is an inexpensive red blend from New York. The makers do not publicize the specific grape used in the wine, but participants generally believed native and hybrid grapes were ingredients. This set it apart from most of the other wines in the set, which largely used traditional vinifera grapes. For fungus, the highest choice proportion was for a Pennsylvania dessert wine (4) made with “vidal blanc” grapes. Although these grapes are probably not in the Taylor, they are hybrids and therefore may be considered somewhat similar. The four other wines with consistent choice proportions for fungus (23, 12, 27, and 35) were also all east coast wines (the only other New York wine in the set is #23). Few wines were regularly chosen to share properties Q or X, but the only other east coast red in the set (35) was highest. Retailers seemed to see some similarity between #14 and the two merlots

(28 and 38), at least with respect to dish and aliens properties. In contrast, connoisseurs were more likely to extend these properties to #29 (an inexpensive California cabernet sauvignon).

Premise wine #19 was the moderately expensive Veuve Clicquot, a French champagne (Table 16). For all properties, the target wine that was chosen most often was #23, the New York state champagne. The two wines, while quite different in terms of region and price, share a distinctive combination of grapes, sparkling quality, and fermentation technique. This was a powerful inductive base for winemakers who picked #23 almost every time for premise #19. Retailers were less likely to choose #23, especially for the fungus property. Other common choices were two non-sparkling French whites (13 and 33). The two other sparkling wines (15 and 37) were from Italy, used different grapes, and had a different taste profile, yielding only moderate selection rates.

The final premise wine, a Gloria Ferrer chardonnay (25), had a very close “sister” (20) that was frequently chosen (though not always). Both were made from the same grapes and by the same producer, but the premise was a reserve and therefore slightly more expensive. There were three other California chardonnays in the set (3, 17, 39) and these were also picked at high rates. Based on the match scores, #17 seems to be a better match than #39, because it is closer in price to the premise ($6.99, $68.99, and $16.99, respectively), but it was picked less often. The difference may be that #17 (a Barefoot chardonnay) is more of a mass market wine whereas #39 and #25 are somewhat more elite wineries. Note that the Louis Jadot Pouilly Fuisse, a French chardonnay (13) was picked fairly often, as was the Ladoucette Pouilly Fume, a French sauvignon blanc (33). Generally, wines that were similar in terms of region had a higher chance of being chosen for the fungus property.

The preceding descriptive analysis suggests that generally, the greater the similarity of the premise and the target the more often the property was extended to the target. The degree of similarity mattered, but so did the combination of different features. The strongest and most consistent extensions were to wines that were high matches on both region and grape varietal. That said, the property definitely mattered: region matches alone were enough to warrant moderate expectations that wines would be affected by the same fungus. Especially for the aliens and dish properties, there was evidence suggesting that varietal mattered and that it could go beyond a literal match of constituent ingredients; there was some consistency in extending properties to certain other grapes. Connoisseurs were willing to extend some properties from a pinot noir to merlots, and all expert types made extensions from a chardonnay to some sauvignon blancs.

These analyses suggest several reasons that the ANOVAs based on the match type scores must be viewed with some caution. First, although the category generation data suggest that the features captured by the match codes are important, they are certainly limited. Certainly other, potentially important properties are not captured by this coding system. Second, it may be that the experimenter's "objective" evaluation of matches does not reflect experts' perceptions, which are what really matters in this case. However, although classifying regional matches by whether they were from the same country makes logical sense and can be consistently applied, it may not have captured the most meaningful clusters for the experts. During their sorts, it was common for experts to begin by making "inconsistent" regional subdivisions, such as: France, California, Eastern US, and Australia/New Zealand (a country, a state, a region spanning states, and a region spanning countries). Similarly, the boundaries for the price categories were set somewhat arbitrarily. However, while limitations to the match type coding scheme exist, I believe that the matches serve a reasonable approximation and are certainly an improvement over examination of straight extension counts.

Similarity and Typicality

Research on similarity and typicality (e.g., Rips, 1975; Gelman & Markman, 1986; Osherson et al., 1990; Sloman, 1993) predicts that the more typical a wine is, the more people will expect other wines to share a property it possesses. Although the concept of typicality is nuanced, it is highly correlated with the similarity of an object to other members of the same category (Rosch & Mervis, 1975; Barsalou, 1985).[42] This leads to several predictions that can be tested with the present data (note, however, that these are exploratory analyses that were not among the main purposes of the present research). To the extent that overall similarity judgments are driving inference, the higher the typicality of a premise wine, the more other wines participants should expect to possess the property. In addition, the more similar a target wine is to the premise wine, the more they should be expected to share an unknown property.

How well did the typicality of a premise wine predict the number of other wines chosen in the inference task? I used the mean category membership ratings (presented in the results section for task 4) as an index of typicality. Another index of wine typicality is its similarity to other wines, calculated here as the average distance between a premise wine and every other wine in the set. For each of the four premise wines, I calculated this score for each expert, based on the full set of sorts they had completed. Table 18 presents the correlations between these two

[42] Researchers have found that how well an object embodies category ideals also influences typicality judgments (Barsalou, 1985; Lynch, Coley, & Medin, 2000; Burnett, Medin, Ross & Blok, 2005).

measures (typicality and average sort distance), the number of wines chosen for each of the five properties in the inference task (for each premise), and the *total* number of wines chosen for each premise.

Table 18. Pearson correlations for indices of typicality and similarity and property extension counts

	Typicality	Avg. sort distance	Q	Xergia	Dish	Fungus	Aliens	Total extension
Typicality	1.000							
Avg. sort distance	-.033	1.000						
Q	.034	-.070	1.000					
Xergia	.351**	-.203*	.035	1.000				
Dish	.207*	-.167	-.015	.447**	1.000			
Fungus	-.006	-.076	.154	.014	-.139	1.000		
Aliens	.553**	-.076	.162	.442**	.661**	-.074	1.000	
Total extension	.342**	-.206*	.496**	.639**	.513**	.537**	.615**	1.000

** Correlation is significant at the 0.01 level (1-tailed)
* Correlation is significant at the 0.05 level (1-tailed)

Individuals' typicality ratings of premise wines (as members of the category "wines") were positively correlated ($p<.05$, 1-tailed) with the number of wines thought to have xergia (.351), to go well with the dish (.207), and to be appreciated by aliens (.553). Typicality was also correlated with the total number of extensions across all property types, but not with "Property Q" or fungus. Average sort distance was not correlated with typicality ratings, and was negatively correlated both with xergia (-.203) and with the total number of extensions across all property types (-.206).

Which of the property types were correlated with each other? The numbers of extensions for each individual property type were all positively correlated with the total number of extensions. Dish, fungus, and X were positively correlated with each other; dish and fungus were highly correlated (.663) and X was moderately correlated with each (dish: .447, fungus: .442). The number of wines thought to possess Property Q was not correlated with the number of extensions for any other property. In sum, the judged typicality of a premise wine, but not its average sort distance, predicted the number of other wines expected to share an unfamiliar property with it.

Justifications

The choices provide some indications about induction strategies, but have some limitations. To examine the results from another angle, I also examined participants' justifications.

Codes and overall frequencies.

I coded the justifications experts used to explain their wine choices with the following eight categories. These categories, which parallel those used to code the wine labels in the sorting tasks, were: (1) Grape, (2) Style (3) Region, (4) Color, (5) Process, (6) Other, (7) Type[43] and (8) Price/Quality. Explanations based on preference were not common for this task, so they were collapsed into the *other* category.

As with the categories for sorting, it was often necessary to combine codes to adequately describe responses. For example, the response, "Same grape, barrel-aged" were coded as Grape

[43] I used *type* for both sparkling and Champagne in this analysis. Because the term Champagne is polysemous—it could refer to grapes, region, technique, or some combination of these—I could have given it multiple codes, but opted to be a bit more conservative here because it was a common component of responses to one of the premise wines.

and Process. Just over half (55%) of the trials could be described by a single code. The rest required two (32%) or three (11%) codes, with just a few trials requiring four or more codes. In all, 771 codes were assigned, an average of 1.6 codes per trial. The code combinations took two different forms. For about one-third of trials (32.6%) the combination represented an intersection: the property was extended to wines that were A and B. Less often (14.6% of trials), the combination was a "split"; the property was extended to wines that were either A or B. For example, 26.7% of the trials that mentioned a grape varietal extended the property to wines that contained any of a set of grapes (e.g., "pinot noirs and chardonnays because that's what's in traditional champagne"). This was distinct from the more typical intersection type of combination (e.g., "only French pinot noir") but parallels an approach observed by Proffitt et al. (2000) among tree experts who, when given an open-ended question with two premises, sometimes "split" their prediction between two genera.

Table 19 depicts the frequency of the code types as a percentage of all codes. Generally, the types of categories that were common for the sorting task were also used in inference. Although the three most common sort category types (color, grape, and region) remained important for inference, their priority shifted. For inference, grape categories were used more than regional ones (the most common category type for sorts). Another shift was the increased use of style-based categories, which moved from sixth-place among sort category types, to the second-most common code for inference justifications. Process-based justifications were also more common for inference than they had been for the sorts.

Table 19. Frequency of Code types

Code	Rank (and %) of Inference Trials	Rank of Sort Categories
Color	4 (16.25)	3
Grape	1 (40.42)	2
Other[44]	6 (15.42)	4
Price/Quality	8 (10.00)	5
Process	5 (16.04)	8
Region	3 (22.92)	1
Style	2 (27.50)	6
Type	7 (11.88)	7

In addition to the category-based codes, I also noted some stylistic patterns. Some justifications (16%) explicitly invoked some sort of causal or ecological explanation in addition to the general content (e.g., "cool climates because the fungus would have to thrive in these frost-prone climates"). Others explicitly used the terms "same" or "similar" (24%).

Justifications by property type

Did the type of property being asked about affect the types of justifications experts used? A two-way ANOVA (Property x Code Type) on the proportion of codes per group found both a significant effect of property, $F(4,80) = 5.622$, p=.00, and a significant property by code interaction, $F(28, 80)$, p =.00.

To understand the substantive nature of the property by code interaction, for each justification code type I compared the means across the five properties. See Figure 16 for a graph

[44] The *other* category is omitted from subsequent figures, but the codes are included in calculations.

of the justifications that showed significant differences among properties. (Means are presented in Appendix X.) Post-hoc Tukey *HSD* tests found the following differences to be significant (*p*s $< .05$). Color and style justifications were used more for the Aliens and Dish properties than for the other three. Price/Quality justifications were used more for Aliens than for Xergia or Fungus, and were also used more for Dish than for Fungus. Grape justifications were used more for Xergia than for Aliens or Dish. Region-based justifications were used more for Fungus than for Aliens, Dish, or Xergia.

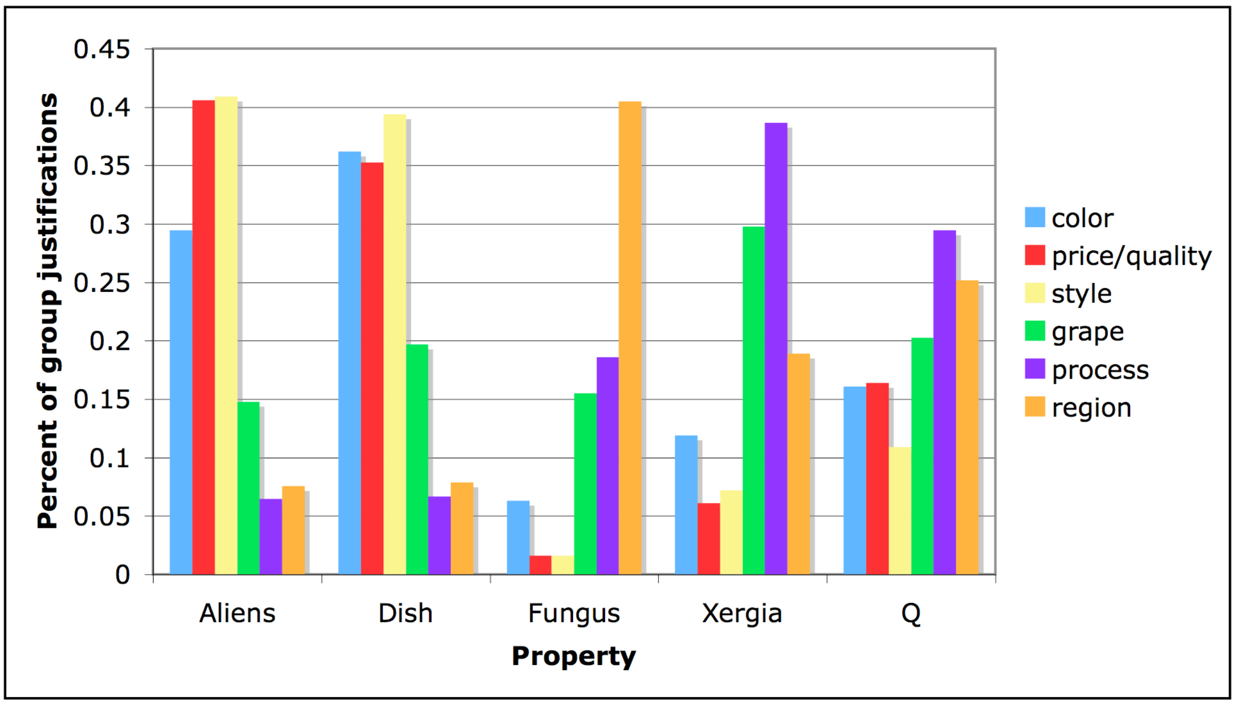


Figure 16. Justification type by property

Note: This graph depicts which types of justifications were used for each type of property for the inference task. The bars represent the mean percent of justifications (by group) of that particular justification type. Two justifications types have been omitted from the figure: *type* (sparkling/still) and *other*. All means are presented in Appendix X.

Group differences in justifications.

Generally, all of the justification types were used by all types of experts. That is, most of the experts in each group invoked each type of explanation at least once. The notable exception was price/quality. Although few winemakers (2 of 9) explained their inferences in terms of price/quality, most connoisseurs (3 of 5) and nearly all retailers (8 of 10) did. Also, although all winemakers and all but one retailer discussed process, fewer connoisseurs (3 of 5) raised the topic.

Table 20. Justification use by group

	Connoisseurs	Retailers	Winemakers
color	0.19	0.15	0.16
grape	0.56	0.24	0.48
other	0.08	0.21	0.13
process	0.10	0.15	0.21
price/quality	0.21	0.12	0.01
region	0.30	0.28	0.13
style	0.22	0.22	0.36
type	0.08	0.13	0.13

Experts did differ, however, in how heavily they relied on the different justification types. The proportion of codes for each trial was calculated for each expert and used as the dependent variable in a mixed 3 x 4 x 5 x 7 ANOVA (Group x Premise x Property x Codetype). A 2-way

ANOVA (Group x Code Type) of code frequency (per expert) showed a significant effect of code type $F(7, 168) = 15.59$, $p = .00$, as well as a significant group by code type interaction, $F(14, 168) = 3.91$, $p = .00$. Post-hoc Tukey's HSD tests revealed that this interaction was driven by retailers using grape-based justifications less than both other groups (ps $< .05$) and by connoisseurs using price/quality justifications more than winemakers ($p < .05$). The effect of group for region was marginally significant, $F(2,21) = 3.18$, $p = .06$. See Table 20 for the pattern of justification use by group.

Summary.

The justification findings suggest that the categories generated in the sorting task were not spurious. Seven of the eight common types of categories from the sorting task were also used in inference justifications. Although all experts used all of these types of categories to some degree, there were a few categories for which there was evidence of differential preference. Retailers used grape-based justifications less than the other two groups. Connoisseurs used price/quality justifications more than winemakers. Both of these differences also were evident in the sorting data (though other differences for the sorting data did not appear for the inference task).

The specific problem—that is, the property in question—also influenced the types of categories emphasized in the justifications. The "Dish" and "Aliens" questions showed a common pattern: more use of color-, style-, and price/quality-based justifications, and less emphasis on region and process. There was moderate use of grape for all properties, but especially for the question about xergia, which also elicited the greatest proportion of justifications addressing process. Justifications about region were rarely used for the dish and aliens questions, appeared moderately in response to xergia and "Property Q," but dominated reasoning about which wines would suffer from the same fungus. In sum, based on the analysis of justifications, the types of categories that had been created in the sorting task also played a role in the reasoning process. There were some differences in which types of justifications the three expert types preferred as well as differences in the types of justifications elicited by questions about different properties.

Discussion

Were the prevailing categories from the sorts also used for inference? It does appear that the category types used most in the sorting task were also frequently used to describe inferences. The relative priority of these types shifted, however, probably in response to the specific probes used in the inference task. For there certainly was an effect of property on type of justification. Style, price/quality, and color were the most important types of justifications for questions about what other wines would go with a novel dish and what other wines aliens would enjoy, but they were not used much when making inferences about what other wines would be affected by a novel fungus, have a novel chemical compound, or possess "Property Q." Justifications invoking regional categories were often used to predict which wines would be likely to have the new fungus, while grape and process-oriented justifications were used for xergia. There was some evidence that different types of experts deployed the categories differently. Retailers were less likely to invoke grape varietal, whereas connoisseurs were more likely to use price/quality categories.

Correlational analyses showed that the rated typicality of a premise wine predicted the number of other wines that would be selected to share a novel property with it. However, similarity—as calculated by the average sorting distance of the premise to other wines—did not. Reversing this analysis to examine the likelihood of a target wine being selected did reveal a relationship between the target's similarity to the premise and its rate of selection. There was a moderate negative correlation, across all property trials: the closer the premise and the target, the more likely the target would be chosen.

There were no differences in these correlations based on expert type, but the sort types (A/B/C) that had emerged in the sorting analyses did seem to be related to the utility of the sorts in predicting target choices for different properties. For the fungus property, the type A (regional) sorts were better predictors. Overall, type B (evaluative) sorts were poor predictors for these properties, but did have weak predictive power for the aliens and dish properties (though weaker than the type A and C sorts). Calculations of the degree to which choices matched the premise on particular dimensions paralleled these findings.

CHAPTER FIVE: GENERAL DISCUSSION

Summary

This research began with the expectation that different types of experts share a knowledge base, but engage in systematically different goal-oriented behaviors. These behaviors were expected to influence the kinds of representations experts formed, leading to differences in the how the groups used categories in reasoning. The findings did not strongly disconfirm these expectations, but neither did they provide powerful support. On the other hand, the domain of wine was a fruitful one for the study of cross-classification.

During the first session, analysis of the activity inventory identified three clusters of experts that broadly conformed to initial group assignments. Factor analysis of their activity ratings revealed sets of behaviors distinctive to each group as well as many shared behaviors.

The original goals of the sorting task were to determine whether experts had access to multiple organizations of the domain, to describe the sorts they created in terms of structure and content, and to identify group differences. The experts did generate multiple sorts with little difficulty, but some characteristics of the sorts were unexpected. Instead of using a single strategy consistently within a sort, experts tended to mix and match categories. This was true at multiple levels of analysis: there was not a strong, consistent sorting profile based on type of expert, individuals used a variety of strategies across sorts, and experts frequently used multiple types of categories within a single sort, as well as multiple types within a given level of a sort. Even at the level of the individual category, there was a striking tendency to mix and match dimensions, as evidenced by the frequency of combination category types, often (but not exclusively) at the initial level of a sort. These combined categories were the second most commonly used category type, and were used by all types of experts. “Pure” sorts—using only a single dimension throughout—were relatively rare. They tended to be fairly shallow in depth and only occurred as a second or third sort, never as the first sort.

However, there were some identifiable patterns in terms of the types of category labels offered by the different groups. Overall, the dominant “pure” category types were region, grape, and color. Note that while the literature on wine expertise has addressed grape (e.g., Hughson, 2003; Solomon, 1990; Solomon, 1997) and color (e.g., Brochet & Dubourdieu, 2001), the influence of region on wine experts’ thinking about wine has not received much attention. Multidimensional scaling of the wines (based on sort-derived distances) across all experts revealed meaningful clusters and two interpretable dimensions. Consistent with the prevailing content category labels, regional clusters were evident, as was a clear division between reds and whites.

Analysis of the sorts showed only weak evidence of differences based on expert type; winemakers used more preference-based category labels than retailers, and retailers used more style-based categories than connoisseurs. Examination of the inter-item distances using the Cultural Consensus Model (Romney et al., 1986) did not reveal strong consensus either overall or by expert type. Neither did multidimensional scaling nor hierarchical cluster analyses of the first sorts yield groups of sorts based on expert type. Instead, the three sort clusters that emerged appear to reflect different approaches: (A) regional, (B) evaluative, and (C) color/grape combinations. It was possible to assign the first sorts to these three categories, but may not be appropriate to assign the experts themselves to the A/B/C groups, as most experts generated more than one sort, and these tended to vary in approach.

Was there evidence that the sort categories were meaningful to the experts in other contexts? The category membership ratings provided support for the legitimacy of the categories tested, as well as confirmatory evidence of cross-classification in the domain: multiple category membership was common. In addition, there was evidence of graded structure for most categories. The results of the similarity task, however, were complicated and did not lend themselves to a clear interpretation. The anticipated priming effects did not appear consistently. Similarity judgments are known to be sensitive to a variety of factors; it is possible that information on the actual wine labels contributed an unintended, confounding prime. Additional experiments and analyses may be able to shed light on these results.

The inference task, on the other hand, revealed some interesting patterns. Although no strong evidence for an influence of expert type emerged, there was evidence that the sorts had an impact on property extension. Averaged across all individuals, there were significant negative correlations between the premise-target distance on the first sort and the likelihood of that target being chosen to share a property with that premise; the closer the two wines, the more likely the selection of the target. Interestingly, when the sorts were broken down by A/B/C type the size and significance of those correlations depended on property type. Evaluative sorts (B) were less predictive overall, but still somewhat useful for mapping extension of the aliens and dish properties. Regional sorts (A), were generally useful, but were strongly correlated with choices for the fungus property.

Another analysis calculated the proportion of selections that were consistent with attention to each of several dimensions (region, color, grape, and price). Grape matches were the highest scoring overall, consistent with Varietal Schema Theory (Hughson, 2005), but as with the correlations, different types of matches were important for different types of properties. Again, region was relatively more useful for inferences about fungus; color and price were somewhat stronger for aliens and dish.

Content analysis of the justifications similarly showed that certain types of justifications were more useful for induction about certain types of properties. For the fungus inference, region-based justifications were most common. For the aliens and dish properties, the dominant justifications drew on color, price/quality, and style. For xergia (and to some extent, property Q), justifications drawing on grape, process, and region were frequent.

The results of this investigation did not find strong occupationally based differences in category organization and use. However, the findings did support the expectation that the domain would be one characterized by cross-classification. It appears that experts in this domain have a variety of ways of organizing and thinking about wines. Although their typical behaviors may have some subtle impact on their preferred approaches, it appears that influence is mild relative to the demands of a particular task.

In general, characterizing experts' sorts was more challenging than anticipated. Giving them the opportunity to create multiple sorts did not lead them to compartmentalize different dimensions, as I had predicted. Instead, they mixed dimensions liberally and somewhat idiosyncratically. This gave the impression of weak consensus, but when asked to judge category membership on a subset of categories, agreement was actually fairly strong.

Implications

Researchers have studied a wide array of categories, varying on a number of dimensions. Some study artificial categories, carefully created and structured by the experimenter whereas others explore real-world categories and concepts. There may be systematic differences in categories based on their content: object categories, folkbiological categories, abstract categories, social categories, mass vs. count categories, relational categories (Medin et al., 2000). Categories have also been categorized based on their themes and relations; while most research has examined taxonomic categories that depend on similarity and class inclusion relationships, other types, such as ad hoc, goal-related categories, and thematic categories such as script-based categories have been identified and studied.

Many of the categories identified and used by the experts in this study appear to be organized around a single feature. Does this make them less valid categories than something like "dog"? In many natural categories, it is true that multiple features are correlated. Yet while the category "red wines" may appear to be based on just a single feature, this feature is embedded in a network of causal and correlational relationships. Wine color results from a particular way of processing grapes and tends to be associated with particular grape varieties. Researchers have found that flavor terms are tightly linked to the color of a wine, to the extent that the flavors attributed to red wines depend highly on other dark or red-colored objects (e.g., red berries, licorice, tar), whereas white wines are described with flavor terms drawing on light-colored objects (e.g., honey, cantaloupe, straw).

Goals of retrieval and use may lead to different organizing systems. Retrieval and use may lead to different ways of organizing items. Most people "file" their clothing taxonomically: pants in one drawer, shirts in another, but use-based systems intermingle: a storage bin of winter outerwear including hats, gloves, snow pants, coats, boots; a timesaving, ready-to-go travel case including toothbrush, toothpaste, shampoo, soap, and small packages of headache tablets. A walk through a grocery store also reflects a mixed organization: on the one hand fresh fruits and vegetables tend to appear in the same area (and sometimes dried fruit, though rarely canned), but you will sometimes find a bin of lemons by the seafood or a rack of bananas in the cereal aisle. Here a combination of storage needs (is refrigeration required or available?), taxonomic groupings, patterns of use, marketing directives, ease of retrieval, customer expectations all seem to interact. Some of these groupings are taxonomic, others might be considered script based, and others perhaps feature-based.

Given earlier findings of expert differences in sorting behavior, why did no strong differences in category structure emerge among these populations? There was good reason to expect that differences in activities and goals would lead the different types of experts to create somewhat different sorts of the wines, perhaps based on utilitarian categories. Research with artificial, experimenter-created categories has shown that differing patterns of category use can affect category representation, leading to differences in feature emphasis (Markman & Ross, 2003). Research on perception identifies ways that the goals of a perceiver can affect where (and whether) category and even feature boundaries are drawn (Goldstone, 1994). There is certainly evidence that how people interact with a domain affects the way they construct categories in it (Boster & Johnson, 1989; Malt, 1995; Medin et al., 1997; Ross, 1997).

Furthermore, several studies have found group based differences in experts' categories (e.g., Lynch et al., 1997; Medin et al. 1997; Medin et al., 2005). However, unlike these folkbiological domains in which group consensus was observed, for wine there does not appear to be any official "scientific taxonomy" that neatly corrals the most important features and dimensions into a definitive hierarchical system of mutually exclusive categories. While there is some consistency in which dimensions are considered important, they are not cleanly correlated with each other. This absence of correlation could lead to even more dramatic differences in groups—in the absence of a "natural" default, the experts' goals and activities could be more powerful drivers of conceptual organization. Or, as appears to be the case here, the cross-cutting nature of these dimensions may lead to even greater overlap among the expert groups.

What might account for this outcome? Note that although there were some behaviors that were distinctive to each group, there were also many shared characteristics. It was not clear that the distinctive interests were sufficient to drive dramatically different organizations of the domain. For example, while a focus on the techniques and process of winemaking is clearly most important to winemakers, it is of interest to consumers as well.

Not only do different types of experts care about similar issues, each expert must take myriad factors into account. Take, for example, this passage from Kevin Zraly's highly esteemed *Complete wine course* (2006):

> What makes one chardonnay different from another? Put it this way: there are many brands of ice cream on the market. They use similar ingredients, but there is only one Ben & Jerry's. The same is true for wine. Among the many things to consider: Is a wine aged in wood or stainless steel? If wood, what type of oak? Was it barrel fermentation? Did the wine undergo a malolactic fermentation? How long does it remain in the barrel (part of the style of the winemaker)? Where do the grapes come from?" (p. 71)

Following this description was a list of recent vintages considered "California Chardonnay Best Bets." As this excerpt illustrates, wine experts—whether connoisseurs, retailers, winemakers, or educators— must take into account a wide range of dimensions: grape, fermentation process, winemakers' styles, region, vintage, and possibly evaluations offered by other experts. It is no wonder that mixed category types were so common.

Open-ended methods such as the free pile sorts used in this dissertation do not constrain the types of similarity judgments being made by subjects. Emphasizing a particular focal dimension (e.g., Chi et al., 1981; Hardiman et al., 1989) may lead to a false impression of consistency and consensus. Certainly, experts could have responded to a request directing them to put together wines that were similar on a particular dimension; in the category membership task they did just that and exhibited reasonable consensus for most categories. Under such instructions, critics might object that the experts were simply selecting wines based on a single arbitrary feature and that these were not meaningful categories. However, the instructions were quite the opposite. Instead of directing experts to put wines together on the basis of a particular dimension, the instructions were to "put together the ones that go together by nature into as many meaningful groups as you like." Experts were not encouraged to create arbitrary groups, but meaningful ones. Is it still possible that the categories were somewhat trivial or ephemeral? Perhaps. It is for this reason that I conducted additional analyses in the second session to determine whether (a subset of the) categories exhibited the type of category structure observed in established categories, and conducted an additional experiment to determine whether the categories were used in an inference task. Certainly these categories may have weaker mental representations than common categories such as "snake" and "bird," yet they are surely at least as well established and coherent as categories that have been studied in research on ad hoc categories (e.g., "things to eat on a diet"), to say nothing of the artificial categories that form the basis of much research on category learning (e.g., Posner & Keele's classic 1968 studies of dot pattern "categories").

The variability of the sorts resulted in a weak fit of the Cultural Consensus Model (Romney et al., 1986). There was neither evidence of distinct, expertise-based sub-groups, nor evidence of strong overall consensus. This type of difficulty is not unprecedented in research on expert sorting behavior. Using the same analytic techniques, Boster and Johnson (1989) found less consistency among expert fisherman than among novices. They theorized that novices' sorts were based largely on the morphological information they could glean from the line drawings that served as stimuli. Because morphological distinctions are highly correlated with taxonomic ones, both novices and experts had access to and used this information in grouping fish. The experts, however, went beyond form and also took function into account. Thus, fish experts' sorts incorporated utilitarian and behavioral groupings as well as taxonomic ones, leading to greater variability. Boster and Johnson concluded that:

> When the development of expertise results in the learning of many alternate devices or bases for structuring a domain, the experts will be more variable in their response than novices and so appear to deviate more often from the consensus. (1989, pp. 882-883)

This interpretation could well apply to wine experts and explain the weak consensus observed.

Wine is different from fish because, among other reasons, morphology is not really an issue. To the naked eye at least, there are not differences in shapes among wines (though bottle shape may be meaningful for experts). At first glance, however, color information is available and researchers in wine expertise have found this to be a quite meaningful characteristic, correlated with the flavor terms, taste expectations, and perhaps even perceptions of a wine

(Brochet & Dubourdieu, 2001). However, there are many other important features of a wine, such as varietal, that are less obvious to the eye, and therefore not so perceptible immediately to a novice. Even when a novice is able to taste a wine, they may not perceive the same features that an expert is able to detect. [45]

The challenge with wine, however, is to make sense of that information. Because the different dimensions are not correlated, I would argue that experts use their knowledge about meaningful intersections of dimensions—combinations of "pure" categories—to drive their thinking about wine. For the expert, it is not just that a combination category is more specific; rather it represents a meaningful group of wines with respect to a particular problem.

Given that a wine has multiple category memberships, how does one decide which is relevant? That, indeed, is one of the challenges for understanding the domain, and it ties back to research on cross-classification. Murphy and Ross (1999, p. 1024) define cross-classification as "when categories overlap in their membership." They further distinguish between a non-competing "inclusion hierarchy" in which the different categories differ in terms of their specificity and "competing categorizations" that have non-overlapping parts as well those that overlap. Take their example comparing an inclusion hierarchy (deer-mammal) and competing categories (deer-meat-driving hazard). Logically, the things that are true of mammals should be true of deer. However, the properties of deer that are shared with prey are different from those it shares with other driving hazards; "the answer as to whether something is edible is quite

[45] On the other hand, wine labels are a means of communication between the purveyor of the wine and consumers of varying levels of expertise. Therefore they may communicate important invisible features more directly. Just as the silhouette of a fish's body (fin configuration, size, visible teeth, etc.) may lead novices to categorize them in ways that are meaningful, so too, may novices be able to glean "deep" information about wines from their labels.

different for deer and driving hazards…Thus it is possible that cross-classification could change the inference one makes about an object" (p. 1025).

The reality is that any item can be categorized in multiple ways: at a minimum there is always the problem of abstraction. Is the object in front of me a plant, a tree, an oak? Even when categories are nested, as in a scientific taxonomy, "which category matters?" is an important question, because it determine how far a property may be extended. Coley, Medin, & Atran (1997) found that even novices with little knowledge about trees had the *expectation* that properties could be extended to other members of the same folk generic.

Patalano, Chin-Parker and Ross (2006) argue that there are several influences on "category preference" when cross-classification makes multiple options available. The literature on social cognition indicates that recently activated categories are influential (e.g., Macrae, Bodenhausen, & Milne, 1995), as are distinctive categories (Nelson & Miller, 1995). Another important issue is the relevance of the categories to the problem (Heit & Rubinstein, 1994; Murphy & Ross, 1999; Kalish & Gelman, 1992; Ross & Murphy, 1999). Rehder (2006) found evidence that for most people, possession of a feature that was considered the "causal antecedent" of the property in question drove choices about extending properties to other items.

Inductive inference is certainly a part of most wine-buying decisions—when buying a sealed bottle, consumers do not have direct information about what is inside. They must make inferences about the contents (and how appropriate they will be for particular needs). This requires the integration of a lot of different inferences. Information on the label, information that you may or may not be able to recall about things mentioned on the label, expectations about what is needed for the situation, the people and food, perhaps, involved. Categories, presumably, allow us to navigate this a bit easier, since likely properties hang together.

Limitations and Future Directions

Instead of considering actual bottles or glasses of wine, experts evaluated wine labels. Although wine labels obviously do not present all of the perceptual features that a bottle of wine does, their ecological validity for this task is fairly high. Many decisions about wine are made using the information summarized on the label. When consumers select wine to purchase, they usually do so without opening the bottle. If a wine has been sampled before, the label (or some other representation of the information presented on the label) is the stimulus most likely used to recollect the experience of drinking the wine. Still, for some of the experts (winemakers especially) it was a bit unnatural to judge the similarity of wines they could not taste.

The small sample size, the restricted geographic range of the experts, and the use of referrals in recruitment, indicate that caution should be taken in generalizing the details of these findings. An obvious example is the fact that winemakers taken from a broader sample would probably drink less east coast wine than those in this study. That said, I would be surprised if the main findings did not generalize. And, because the collection and analysis of this kind of data was extremely time-consuming, I would hesitate to simply "ramp up" the study. However, using internet-based data collection might be one practical way to expand the size and geographic range of the sample.

The small sample size also imposed a procedural limitation; it was preferable for all experts to complete all tasks, in the same order. It is possible that the obvious emphasis on categories in the sorting tasks of the first session and the category membership judgments of the second session may have influenced subsequent inference justifications. Again, a larger pool of subjects would allow some flexibility in what each individual needs to complete, reducing the risk of inter-task contamination.

In the sorting task, there were few significant group differences in the use of different types of categories. However, the means sometimes hinted at other differences that were swamped by within-group variability. A larger sample size might reveal more stable differences. In addition, analysis of responses on the feature listing component of the card review task could provide another perspective on group differences in emphasis.

Obviously it would be interesting to obtain a more complete set of novice data. Much of the information the experts used was visible on the labels—would novices having access to this same data create comparable sorts? In particular, would the combination categories so prevalent among experts emerge for novices as well?

Analysis of the information visible on the labels might also help explain the puzzling results of the similarity task. It is possible that this information was contributing an additional prime, independent of the intended manipulation. Systematic investigation of the content of the physical labels could help inform this analysis.

A potentially interesting extension of this research would be a network analysis of the experts to examine the relationships among them and the degree to which attitudes and patterns of category use are transmitted through social contact. Though well populated, the world of wine has some well-traveled byways, leading to a potentially large amount of contact and information transmission among inhabitants.

Conclusion

Despite an understandable, persistent assumption in research on concepts and cognition that concepts are organized in neat taxonomic hierarchies, this is an incomplete portrayal of the true nature of concepts. This dissertation adds to a growing literature showing that there is great complexity in concepts. In particular, this research demonstrates the prevalence of cross-classification in some domains, which poses a challenge for models of concepts, knowledge, memory, and reasoning. The field should strive to develop models and theories that account for a broader range of phenomena. This dissertation documents and begins to explain one set of challenges. In addition, it documents the effects of extended, goal-directed interaction with a domain on conceptual structure and use, in a real world setting. Future work will explore these issues further, hopefully taking these phenomena and generating a model that can explain and predict their formation and impact.

REFERENCES

Ahn, W. K. (1998). Why are different features central for natural kinds and artifacts?: The role of causal status in determining feature centrality. *Cognition, 69*, 135-178.

Ahn, W. K., & Kim, N. S. (2001). The causal status effect in categorization - An overview. In *Psychology of Learning and Motivation: Advances in Research and Theory,* (Vol. 40, pp. 23-65).

Anderson, J. R. (1987). Skill acquisition: Compilation of weak method problem solutions. *Psychological Review, 12*(1), 84-89.

Bailenson, J. N., Shum, M. S., Atran, S., Medin, D. L., & Coley, J. D. (2002). A bird's eye view: biological categorization and reasoning within and across cultures. *Cognition, 84*(1), 1-53.

Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition, 10*(1), 82-93.

Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition, 11*(3), 211-227.

Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 629-654.

Barsalou, L. W. (1991). Deriving categories to achieve goals. *Psychology of Learning and Motivation-Advances in Research and Theory, 27*, 1-64.

Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution.* Berkeley, CA: University of California Press.

Boster, J. S., & Johnson, J. C. (1989). Form or function: A comparison of expert and novice judgments of similarities among fish. *American Anthropologist, 91*(4), 866-889.

Brochet, F. (2001). Chemical object representation in the field of consciousness. Prix coup de coeur, Academie Amorin, France.

Brochet, F., & Dubourdieu, D. (2001). Wine descriptive language supports cognitive specificity of chemical senses. *Brain and Language, 77*(2), 187-196.

Burnett, R. C., Medin, D. L., Ross, N. O., & Blok, S. V. (2005). Ideal is typical. *Canadian Journal of Experimental Psychology, 59*(1), 3-10.

Chase, W. G. & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*(1), 55-81.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and Representation of Physics Problems by Experts and Novices. *Cognitive Science, 5*(2), 121-152.

Coley, J. D., Medin, D. L., & Atran, S. (1997). Does rank have its privilege? Inductive inferences within folkbiological taxonomies. *Cognition, 64,* 73-112.

Collins, A. M. & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8*, 241-248.

Dreyfus, H. L. (1997). Intuitive, deliberative, and calculative models of expert performance. In C.E. Zsambok & G. Klein (Eds.), *Naturalistic decision making* (pp. 17-28). Hillsdale, NJ: Lawrence Erlbaum Associates.

Ericsson, K. A., & Charness, N. (1994). Expert Performance - Its Structure and Acquisition. *American Psychologist, 49*(8), 725-747.

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.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology, (47)*, 273-305.

Ericsson, K. A, & Staszewski, J. J. (1989). Skilled memory and expertise: Mechanisms of exceptional performance. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon*, (pp. 235-267). Hillsdale, NJ: Lawrence Erlbaum.

Fiske, S. T. (1998). Stereotyping, prejudice and discrimination. In D. T. Gilbert, S. T. Fiske, and G. Lindzey (Eds.), *Handbook of social psychology,* (pp. 357-411). New York: McGraw-Hill.

Garson, G. David (2008). "Multi-dimensional scaling", from *Statnotes: Topics in Multivariate Analysis.* Retrieved 05/08/2008 from http://www2.chass.ncsu.edu/garson/pa765/statnote.htm.

Gelman, S. A., Coley, J., & Gottfried, G. M. (1994). Essentialist beliefs in children: The acquisition of concepts and theories. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the Mind: Domain Specificity in Cognition and Culture* (pp. 341-366). Cambridge: Cambridge University Press.

Gelman, S. A., & Markman, E. M. (1986). Categories and Induction in Young Children. *Cognition, 23*, 183-209.

Glaser, R., & Bassok, M. (1989). Learning theory and the study of instruction. *Annual Review of Psychology, 40*, 631-666.

Gobbo, C., & Chi, M. (1986). How Knowledge Is Structured and Used by Expert and Novice Children. *Cognitive Development, 1*(3), 221-237.

Gobet, F. (1998). Expert memory: a comparison of four theories. *Cognition, 66*, 115-152.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 49*, 585-612.

Hardiman, P. T., Dufresne, R., & Mestre, J. P. (1989). The relation between problem categorization and problem-solving among experts and novices. *Memory & Cognition, 17*(5), 627-638.

Hassebrock, F., Johnson, P. E., Bullemer, P., Fox, P. W., & Moller, J. H. (1993). When Less Is More - Representation and Selective Memory in Expert Problem-Solving. *American Journal of Psychology, 106*(2), 155-189.

Heit, E., & Rubinstein, J. (1994). Similarity and Property Effects in Inductive Reasoning. *Journal of Experimental Psychology, 20*(2), 411-422.

Hughson, A. L. (2003). *Cork and talk: The cognitive and perceptual bases of wine expertise.* Unpublished doctoral dissertation, University of Sydney, Sydney.

Hughson, A. L., & Boakes, R. A. (2001). Perceptual and cognitive aspects of wine expertise. *Australian Journal of Psychology, 53*(2), 103-108.

Inhelder, B. & Piaget, J. (1964). *The Early Growth of Logic in the Child: Classification and Seriation*. London: Routledge and Kegan Paul.

Jee, B. D., & Wiley, J. (2004). The Influence of Goal-directed Activity on Categorization and Reasoning. *Working Paper*.

Johnson, K. E., & Mervis, C. B. (1997). Effects of varying levels of expertise on the basic level of categorization. *Journal of Experimental Psychology-General, 126*(3), 248-277.

Kalish, C. W., & Gelman, S. A. (1992). On wooden pillows: Multiple classification and children's category-based inductions. *Child Development, 63*, 1536-1557.

Klein, G. (1997). Developing expertise in decision making. *Thinking & Reasoning, 3*(4), 337-352.

Komatsu, L. K. (1992). Recent Views of Conceptual Structure. *Psychological Bulletin, 112*(3), 500-526.

Kunda, Z, & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review, 103*(2), 284-308.

Lehrer, A. (1975). Talking About Wine. *Language, 51*(4), 901-923.

Lehrer, A. (1983). *Wine and conversation*. Bloomington: Indiana University Press.

Lesgold, A., Rubinson, H., Feltovich, P., Glaser, R., Klopfer, D., Wang, Y. (1988). Expertise in a complex skill: diagnosing x-ray pictures. In: M. T. Chi, R. Glaser, M. Farrar (Eds.) *The nature of expertise.* (pp. 311–341). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Lin, E. L. (1997). Thematic relations in adults' concepts and categorization. (Doctoral dissertation, University of Illinois, Urbana-Champaign, IL, 1997). *Dissertation Abstracts International, 57*, 5372.

Lin, E. L., & Murphy, G. L. (2001). Thematic relations in adults' concepts. *Journal of Experimental Psychology: General, 130*, 3-28.

Lingle, J. H., Altom, M. W., & Medin, D. L. (1984). Of cabbages and kings: Assessing the extendibility of natural object concepts to social things. In R. S. Wyer & T. K. Srull (Eds.) *Handbook of social cognition* (Vol. 1, pp. 71-118). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Lopez, A., Atran, S., Coley, J. D., Medin, D. L., & Smith, E. E. (1997). The tree of life: Universal and cultural features of folkbiological taxonomies and inductions. *Cognitive Psychology, 32*, 251-295.

Love, B. C. (2005). Environment and goals jointly direct category acquisition. *Current Directions in Psychological Science, 14*(4), 195-199.

Lynch, E. B., Coley, J. B., & Medin, D. L. (2000). Tall is typical: Central tendency, ideal dimensions, and graded category structure among tree experts and novices. *Memory & Cognition, 28*, 41-50.

Macrae, C. N., & Bodenhausen, G. V. (2000). Social Cognition: Thinking Categorically about Others. *Annual Review of Psychology, 51*, 93-120.

Malt, B. C. (1995). Category coherence in cross-cultural perspective. *Cognitive Psychology, 29*, 85-148.

Markman, A. B., & Ross, B. H. (2003). Category Use and Category Learning. *Psychological Bulletin, 129*(4), 592-613.

Markman, A. B., Yamauchi, T., & Makin, V. S. (1997). The creation of new concepts: A multifaceted approach to category learning. In T. B. Ward, S. M. Smith & J. Vaid (Eds.), *Creative thought: An investigation of conceptual structures and processes*. Washington, D.C.: American Psychological Association.

Markman, E. M. (1989). *Categorization and Naming in Children: Problems of Induction.* Cambridge, MA: MIT Press.

McCoy, E. (2005). *The Emperor of Wine*. New York: HarperCollins.

Medin, D. L., Coley, J. D., Storms, G., & Hayes, B. K. (2003). A relevance theory of induction. *Psychonomic Bulletin & Review, 10*(3), 517-532.

Medin, D. L., & Goldstone, R. L. (1995). The predicates of similarity. In C. Cacciari (Ed.), *Similarity in Language, Thought, and Perception*. (pp. 83-110). Brussels: BREPOL.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100,* 254-278.

Medin, D. L., Lynch, E. B., Coley, J. D., & Atran, S. (1997). Categorization and reasoning among tree experts: Do All roads lead to Rome? *Cognitive Psychology, 32*(1), 49-96.

Medin, D. L., Lynch, E. B., & Solomon, K. O. (2000). Are there kinds of concepts? *Annual Review of Psychology, 51*, 121-147.

Medin, D. L., & Rips, L. J. (2005). Concepts and categories: meaning, memory, and metaphysics. In K. J. Holyoak & R. G. Morrison (Eds.), *Cambridge handbook of thinking and reasoning* (pp. 37-72). Cambridge, England: Cambridge University Press.

Medin, D. L., Ross, N. O., Atran, S., Cox, D., Coley, J., Proffitt, J. B., et al. (2005). Folkbiology of freshwater fish. *Cognition, 99*(3), 237-273.

Melcher, J. M., & Schooler, J. S. (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.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior 16*(5), 519-533.

Morrot, G., Brochet, F., & Dubourdieu, D. (2001). The Color of Odors. *Brain and Language.*

Murphy, G. L. (2001). Causes of taxonomic sorting by adults: A test of the thematic-to-taxonomic shift. *Psychonomic Bulletin & Review, 8*(4), 834-839.

Murphy, G. L. (2004). *The big book of concepts* (1st MIT Press paperback ed.). Cambridge, Mass.: MIT Press.

Murphy, G. L. & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*, 289-316.

Murphy, G. L. & Ross, B. H. (1999). Induction with cross-classified categories. *Memory & Cognition, 27*(6), 1024-1041.

Murphy, G. L., & Wright, J. C. (1984). Changes in conceptual structure with expertise:

Differences between real-world experts and novices. *Journal of Experimental Psychology-Learning Memory and Cognition, 10*(1), 144-155.

Neisser, U. (1987). From direct perception to conceptual structure. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 11-24). Cambridge: Cambridge University Press.

Nelson, L. J., & Miller, D. T. (1995). The distinctiveness effect in social categorization: You are what makes you unusual. *Psychological Science, 6*(4), 246-249.

Nguyen, S. P. (2007). Cross-classification and category representation in children's concepts. *Developmental Psychology, 43*(3), 719-731.

Noble, A. C., Arnold, R. A., Buechsenstein, J., Leach, E. J., Schmidt, J. O., & Stern, P. M. (1987). Modification of a standardized system of wine aroma terminology. *American Journal of Enology and Viticulture, 38*(2), 143-146.

Norman, G. R., Brooks, L. R., & Allen, S. W. (1989). Recall by expert medical practitioners and novices as a record of processing attention. *Journal of Experimental Psychology-Learning Memory and Cognition, 15*(6), 1166-1174.

Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review, 97*, 185-200.

Palmer, C. F., Jones, R. K., Hennessy, B. L., Unze, M. G., & Pick, A. D. (1989). How is a trumpet known? The "basic object level" concept and the perception of musical instruments. *American Journal of Psychology, 102*, 17-37.

Pangborn, R., Berg, H., & Hansen, B. (1963). The influence of color on discrimination of sweetness in dry table-wine. *American Journal of Psychology, 76*, 492-495.

Patalano, A., Chin-Parker, S., & Ross, B. (2006). The importance of being coherent: Category coherence, cross-classification, and reasoning. *Journal of Memory and Language, 54*, 407-424.

Patel, V. L., Arocha, J. F., & Kaufman, D. R. (1994). Diagnostic reasoning and medical expertise. *Psychology of Learning and Motivation, 31*, 187-252.

Proffitt, J. B., Coley, J. D., & Medin, D. L. (2000). Expertise and category-based induction. *Journal of Experimental Psychology-Learning Memory and Cognition, 26*(4), 811-828.

Posner, M. I. & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77*, 353-363.

Rabin, M. D., & Cain, W. S. (1984). Odor recognition: Familiarity, identifiability, and encoding consistency. *Journal of Experimental Psychology-Learning Memory and Cognition,* *10*(2), 316-325.

Ratneshwar, S., Barsalou, L. W., Pechmann, C., & Moore, M. (2001). Goal-derived categories: The role of personal and situational goals in category representations. *Journal of Consumer Psychology, 10*(3), 147-157.

Rehder, B. (2006). When similarity and causality compete in category-based property generalization. *Memory & Cognition, 34*(1), 3-16.

Rips, L. J. (1975). Inductive judgments about natural categories. *Journal of Verbal Learning and Verbal Behavior, 14*(6), 665-681.

Rips, L. J. (2001). Necessity and natural categories. *Psychological Bulletin, 127*(6), 827-852.

Rips, L. J., & Estin, P. A. (1998). Components of objects and events. *Journal of Memory and Language*, *39*, 309-330.

Romney, A. K., Weller, S. C., & Batchelder, W. H. (1986). Culture as consensus: A theory of culture and informant accuracy. *American Anthropologist, 88*, 313-338.

Rosch, E., & Mervis, C. B. (1975). Family resemblances - studies in internal structure of categories. *Cognitive Psychology, 7*(4), 573-605.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic Objects in Natural Categories. *Cognitive Psychology, 8*(3), 382-439.

Ross, B. H. (1997). The use of categories affects classification. *Journal of Memory and Language, 37*, 240-267.

Ross, B. H., & Murphy, G. L. (1999). Food for thought: Cross-classification and category organization in a complex real-world domain. *Cognitive Psychology, 38*, 495-553.

Ross, J. P. (2002, Winter). Psychology of Quality. *Global Vintage Quarterly*.

Schank, R. C., Collins, G. C., & Hunter, L. E. (1986). Transcending inductive category formation in learning. *The Behavioral and Brain Sciences, 9*(4), 639-686.

Schoenfeld, A. H. & Hermann, D. J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, 484-494.

Shafto, P., & Coley, J. (2003). Development of categorization and reasoning in the natural world: Novices to experts, naive similarity to ecological knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(4), 641-649.

Shafto, P., Kemp, C., Baraff, E., Coley, J., & Tenenbaum, J. (2005). Context-sensitive induction.

In the *Proceedings of the Twenty-Seventh Annual Meeting of the Cognitive Science Society*.

Shanteau, J. (1992). How much information does an expert use? Is it relevant? *Acta Psychologica, 81*, 75-86.

Shiffrin, R. M. (1996). Laboratory experimentation on the genesis of expertise. In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games*. Mahwah: Lawrence Erlbaum.

Sloman, S. A. (1993). Feature-Based Induction. *Cognitive Psychology, 25*, 231-280.

Sloman, S. A. (1998). Categorical Inference is not a Tree: The Myth of Inheritance Hierarchies. *Cognitive Psychology, 35*, 1-33.

Sloman, S. A., Love, B. C., & Ahn, W. (1998). Feature centrality and conceptual coherence. *Cognitive Science, 22,* 189-228.

Sloutsky, V. M. (2003). The role of similarity in the development of categorization. *TRENDS in Cognitive Science, 7*(6), 246-251.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and Process in Semantic Memory - Featural Model for Semantic Decisions. *Psychological Review, 81*(3), 214-241.

Solomon, G. E. A. (1990). Psychology of Novice and Expert Wine Talk. *American Journal of Psychology, 103*(4), 495-517.

Solomon, G. E. A. (1997). Conceptual change and wine expertise. *Journal of the Learning Sciences, 6*(1), 41-60.

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise - Is the basic level in the eye of the beholder. *Cognitive Psychology, 23*(3), 457-482.

Tversky, B. & Hemenway, K. (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General, 113*(2), 169-193.

Weller, S. C. (2007). Cultural consensus theory: Applications and frequently asked questions. *Field Methods, 19*(4), 339-368.

Zraly, K. (2005). *Windows on the world complete wine course.* New York: Sterling Publishing.

APPENDIX A:

THE WINE KNOWLEDGE TEST AND RESULTS

Questions			C	R	W	Expert Total	Novices
1. Indicate the traditional color of the following varieties of wine:							
Chardonnay	**White**	Red				100%	73%
Shiraz	White	**Red**				100%	73%
Merlot	White	**Red**				100%	100%
Chambourcin	White	**Red**	100%	75%	91%	87%	73%
Riesling	**White**	Red				100%	100%
Semillon	**White**	Red				100%	55%
Gewürztraminer	**White**	Red				100%	73%
Grenache	White	**Red**	100%	100%	91%	97%	27%
2. How do Botrytis wines differ from standard wines? A. Sugar is added to standard still wine to increase sweetness **B. Grapes are infected by a fungus called Botrytis** C. Grapes of the variety Botrytis are used D. Botrytis fermentation techniques are used E. None of the above			100%	83%	100%	93%	9%

3. What is the main grape variety used in a "Chianti"? A. Semillon **B. Sangiovese** C. Cabernet D. Nebbiolo E. Pinot Noir	86%	92%	91%	90%	27%
4. What is the distinction between aroma and bouquet? A. Bouquet is produced by red grapes and aroma by white grapes B. Bouquet occurs only in sparkling wines and aroma occurs only in still wines C. Aroma is based on climate, bouquet on soils **D. Bouquet comes from fermentation procedures whereas aroma comes from the grape** E. Bouquet fades with bottle age whereas aroma does not	43%	50%	73%	57%	54%
5. Which grapes are used to make traditional champagne? A. Riesling and Chardonnay B. Shiraz and Cabernet **C. Chardonnay and Pinot Noir** D. Grenache and Semillon E. Sauvignon Blanc				100%	0%

6. What color is the flesh of a Pinot Noir grape? A. Red **B. White** C. Pink D. Purple E. Yellow	43%	75%	82%	70%	18%
Percent correct	90%	90%	94%	92%	52%

Note. Percentages are percent correct. Correct responses are in bold.

Where performance varied by expert group, group breakdowns appear. Performance scores are out of 13 (not 6) points because the eight sub-parts of first item count separately.

Item # 4 (aroma vs. bouquet) was challenging; several experts complained that none of the answers was correct. There may be a cultural difference between the US and Australian use of these terms.

APPENDIX B:

BACKGROUND QUESTIONNAIRE

A1. How often do you drink wine? (*circle one*)

a. At least once a day
b. At least once a week
c. At least once a month
d. At least once a year
e. Never

A2. How large is your personal wine cellar? ☐

A3. How many different wines would you estimate you taste a year? (*circle one*)

a. More than 1000
b. 501-1000
c. 251-500
d. 126-250
e. 51-125
f. 26-50
g. 11-25
h. 10 or fewer

A4. How many years have you been tasting wine at about that rate? ☐

A5. How many different wines would you estimate that you have tasted in your lifetime?

a. More than 10,000
b. 5001-10,000
c. 2501-5000
d. 1251-2500
e. 501-1250
f. 251-500
g. 101-250
h. 51-100
i. 26-50
j. 11-25
k. 10 or fewer

A6. Of the past 100 bottles of wine you have tasted, estimate the percentage that came from each of the following regions. If you have not tasted 100 wines, base your percentage on the number you have tasted.

a. ☐ California
b. ☐ Oregon
c. ☐ Washington
d. ☐ New York
e. ☐ Pennsylvania
f. ☐ France
g. ☐ Italy
h. ☐ Germany
i. ☐ Spain
j. ☐ Australia
k. ☐ Chile
l. ☐ Argentina
m. ☐ Other (please specify): ____________________

B1. How would you characterize your past and present professional involvement in the wine industry? (*circle all that apply*):

a. No professional activity
b. Commercial wine sales or marketing professional
c. Wine retail (owner or employee)
d. Sommelier or wine steward
e. Wine educator
f. Winemaker
g. Wine scientist or researcher
h. Wine writer
i. Other wine-related professional activity. If so, please specify:

B2. How would you characterize your personal involvement with wine? (*circle all that apply*):

a. I am a collector of wine.
b. I participate in a wine tasting club with friends.
c. My family usually served wine with dinner when I was a child.
d. Wine is not particularly important to me.
e. None of the above.

B3. How much have you read about wine? (*circle one*)

a. 3 or more books or articles
b. 1-3 books or articles
c. Less than 1 book or article
d. Only labels

B4. What wine-related jobs have you had, if any? Please list each job and its duration.

B5. What wine organizations are you a member of? (*circle all that apply*)

a. American Society for Enology and Viticulture
b. American Wine Society
c. Dionysian Society
d. Pennsylvania Wine Society
e. Pennsylvania Wine Association
f. Society of Wine Educators
g. The Wine Brats
h. Wine America (The National Association of American Wineries)
i. Wine America Trailblazers
j. Others. If so, please specify:

B6. Please describe your <u>wine-related</u> education (*circle all that apply*)

a. Self-taught
b. Reading

c. Non-academic courses for professional certification.
   If so, how many? ☐
d. Other non-academic, "Serious" courses (exams).
   If so, how many? ☐
e. "Fun" courses (no exams).
   If so, how many? ☐
f. Academic degree. If so, please specify the institution(s), the degree(s) and the major(s) or concentration(s):

B7. A number of different certifications exist in the wine industry, such as the Certified Specialist of Wine, the Certified Wine Educator designation, the Masters of Wine, the Wine & Spirit Education Trust Awards, the Culinary Institute of America Certifications, AWS Wine Judge Certification, and the Court of Master Sommeliers' MS. If you have any of these or other comparable qualifications, please list them below. Be as specific as possible.

B8. How often do you engage in the following activities? (*Responses options were: Never, Rarely, Occasionally, Often, and Regularly)*

	Activity
B8-1.	Score a wine using a structured evaluation.
B8-2.	Give a verbal description of a wine.
B8-3.	Taste a wine "blind".
B8-4.	Write about wine for the general public.
B8-5.	Write about wine for a specialized audience.
B8-6.	Speak about wine to groups of ten or more.
B8-7.	Speak about wine to novices.
B8-8.	Speak about wine to a knowledgeable audience.
B8-9.	Plant or tend grapevines.
B8-10.	Harvest grapes.
B8-11.	Taste grapes.
B8-12.	Study wine sales patterns.

B8-13.	Study wine rankings and/or ratings.
B8-14.	Serve wine to customers.
B8-15.	Set wine prices.
B8-16.	Think about wine-food pairings for self or family.
B8-17.	Recommend wine-food pairings to strangers.
B8-18.	Make decisions about when to harvest grapes.
B8-19.	Make decisions about type of barrel to use.
B8-20.	Sell wine by the glass.
B8-21.	Sell wine by the bottle.
B8-22.	Sell wine by the case.
B8-23.	Sell wine by the barrel.
B8-24.	Taste wine for flaws.
B8-25.	Put together a wine list.
B8-26.	Tour wineries.
B8-27.	Help others choose which wine to buy.
B8-28.	Encourage others to buy a particular wine.
B8-29.	Describe wine-making techniques to others.
B8-30.	Think about disease prevention for grapes.
B8-31.	Attend wine tastings.
B8-32.	Conduct wine tastings.
B8-33.	Think about pest management.
B8-34.	Sample wines before they are mature.
B8-35.	Use heavy machinery.

B8-36.	Research new wine-making procedures.
B8-37.	Manage inventory.
~~B8-38.~~	~~Demonstrate wine expertise.~~
B8-39.	Plan special events to promote a wine.
B8-40.	Conduct a sensory evaluation of a wine.
B8-41.	Conduct a chemical evaluation of a wine.

C1. What year were you born? [ ]

C2. Are you: A. Male B. Female

C3. What is your current profession?
  a. Title: ____________________
  b. How long have you had this position? [ ]

C4. What is the highest level of education you reached? (*circle one*)
  a. Less than high school equivalent
  b. High school
  c. Some college
  d. College degree
    Major(s): ____________________
  e. Some graduate study
    Field(s): ____________________
  f. Graduate degree(s)
    Field(s): ____________________
    Degree(s):__________________

APPENDIX C:

BEHAVIORAL SELF-REPORTS FROM BACKGROUND QUESTIONNAIRE

The data in Table C1 shows that the tasting range for connoisseurs was fairly narrow (51-500 wines annually), whereas for winemakers the pattern was bimodal. Seven fell in the same 51-500 wine range as the connoisseurs, but another three reported tasting more than 1000 wines per year. See Tables C2, C3, and C4 for additional information about the experts' behavior.

Table C1. Number of different wines tasted annually by expert type (Item A3)

Group	11-25	26-50	51-125	126-250	251-500	501-1000	>1000	Total
Connoisseurs	1		2	2	2			7
Retailers		1		1	2	3	5	12
Winemakers			4	2	1	1	3	11
Total	1	1	6	5	5	4	8	30

Table C2. Total number of wines tasted by expert type (Item A5)

Group	101-250	251-500	501-1250	1251-2500	2501-5000	5001-10,000	10,000	Total
Connoisseurs		1	1	2	2	1		7
Retailers	1				2	3	6	12
Winemakers				4	3	1	3	11
Total	1	1	1	6	7	5	9	30

The following pie charts depict group averages based on experts' estimates of the regional distribution of the past 100 bottles of wine they had tasted over 13 regional categories.

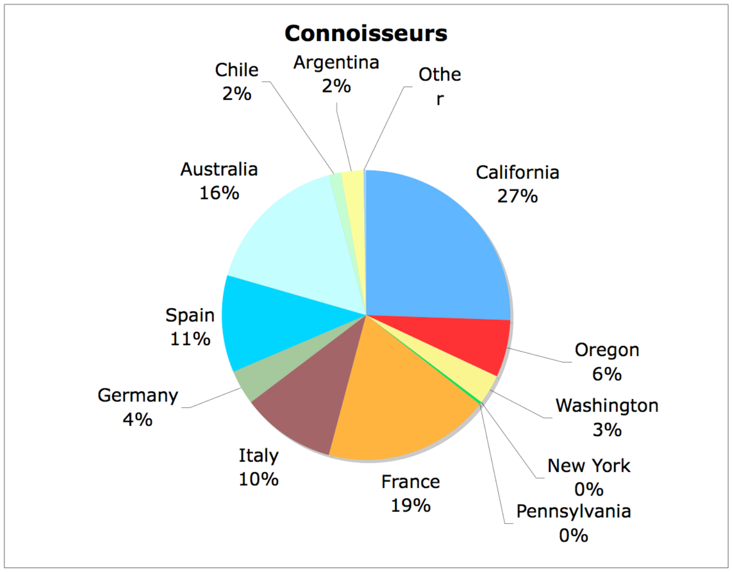


Figure C1. Connoisseur wine consumption by region

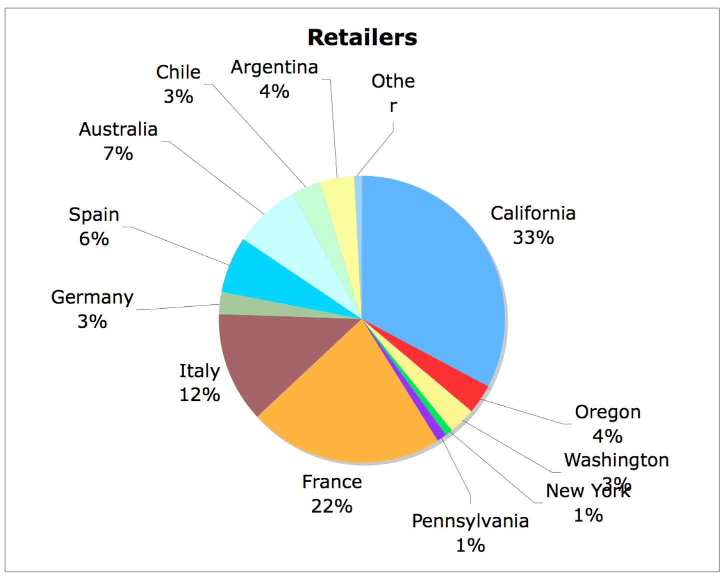


Figure C2. Retailer wine consumption by region

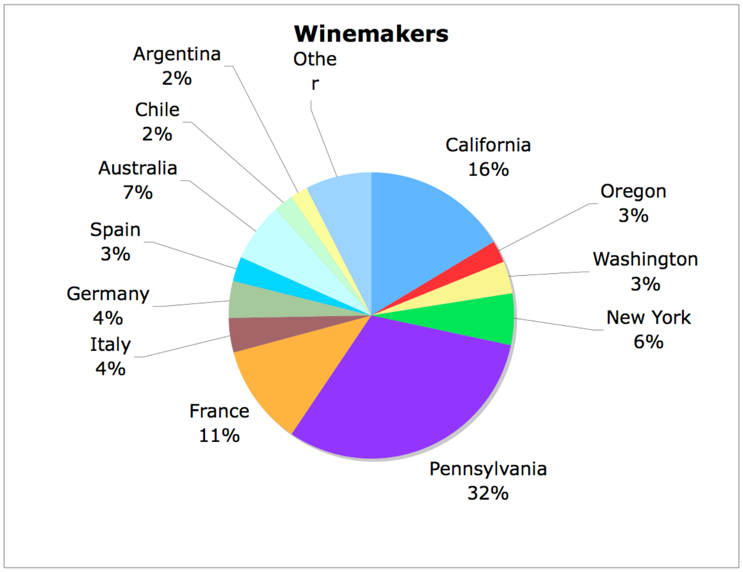


Figure C3. Winemaker wine consumption by region

Table C3. Characteristics of personal involvement with wine (Item B2)

	Connoisseurs (n = 7)	Retailers (n = 12)	Winemakers (n = 11)	Total
A wine collector	5	5	6	16
In a tasting club	3	4	8	15
Drank wine at family dinners		3	3	6
Wine is not important to me		1		1
None of above[46]	1	3	2	6

[46] The “none of the above” option was added early in data collection, prompted by a number of respondents who felt that none of the other options described their involvement with wine well.

Table C4. Wine related education (Item B6)

	Connoisseurs (n = 7)	Retailers (n = 12)	Winemakers (n = 11)	Total
Self-taught	6	11	7	24
Reading	7	10	8	25
Non-academic courses for certification	1	4	4	9
Other "serious" Non-academic courses	1	2	3	6
"Fun" Non-academic courses	5	3	3	11
Academic		3	7	10

Table C5. Professional involvement with wine (Item B1)

	Connoisseurs (n = 7)	Retailers (n = 12)	Winemakers (n = 11)	Total
No professional activity	4			4
Commercial wine sales or marketing		5	2	7
Wine retail (owner or employee)		8	6	14
Sommelier or wine steward		5		5
Wine educator	2	5	5	12
Winemaker		1	11	12
Wine scientist or wine researcher			2	2
Wine writer	1	3	2	6
Other professional activity	2	3	5	10

APPENDIX D:

BACKGROUND QUESTIONNAIRE ACTIVITY INVENTORY

Item text	Statistic	Revised Groups			
		C	R	W	Total
Score a wine using a structured evaluation	Mean	3.29	2.50	3.73	3.13
	Std. Dev.	1.496	1.000	1.348	1.332
Give a verbal description of a wine	Mean	3.86	4.42	4.64	4.37
	Std. Dev.	.900	.900	.674	.850
Taste a wine "blind"	Mean	2.71	3.50	4.27	3.60
	Std. Dev.	1.254	1.087	.905	1.192
Write about wine for general public	Mean	1.43	2.33	2.73	2.27
	Std. Dev.	.535	1.303	1.679	1.388
Write about wine for specialized audience	Mean	1.86	2.50	3.27	2.63
	Std. Dev.	1.464	1.243	1.348	1.402
Speak about wine to groups (10+)	Mean	2.14	3.58	3.73	3.30
	Std. Dev.	1.464	1.240	.905	1.317
Speak about wine to novices	Mean	2.57	4.08	4.09	3.73
	Std. Dev.	1.512	1.084	1.300	1.388
Speak about wine to knowledgeable audience	Mean	2.14	3.42	3.55	3.17
	Std. Dev.	1.215	1.240	.934	1.234
Plant or tend grapevines	Mean	1.71	1.25	3.45	2.17
	Std. Dev.	1.496	.622	1.635	1.599
Harvest grapes	Mean	1.71	1.42	3.45	2.23
	Std. Dev.	1.496	.515	1.635	1.547
Taste grapes	Mean	2.14	1.83	4.55	2.90
	Std. Dev.	1.464	.835	1.214	1.689
Study wine sales patterns	Mean	1.86	3.25	4.36	3.33
	Std. Dev.	1.464	1.603	.674	1.583
Study wine rankings/ratings	Mean	4.00	3.92	2.91	3.57
	Std. Dev.	1.414	.996	1.221	1.251
Serve wine to customers	Mean	1.00	2.83	3.91	2.80
	Std. Dev.	.000	1.403	1.375	1.627

Set wine prices	Mean	1.00	3.67	3.73	3.07
	Std. Dev.	.000	1.826	1.618	1.874
Think about wine-food pairings for self or family	Mean	4.71	4.42	4.64	4.57
	Std. Dev.	.488	.900	.674	.728
Recommend wine-food pairings to strangers	Mean	3.29	4.75	4.27	4.23
	Std. Dev.	1.380	.452	.905	1.040
Decide when to harvest grapes	Mean	1.29	1.00	4.45	2.33
	Std. Dev.	.488	.000	1.293	1.826
Decide what barrel type to use	Mean	1.29	1.00	4.55	2.37
	Std. Dev.	.756	.000	1.214	1.866
Sell wine by glass	Mean	1.00	2.33	2.55	2.10
	Std. Dev.	.000	1.614	1.635	1.517
Sell wine by bottle	Mean	1.00	4.08	3.73	3.23
	Std. Dev.	.000	1.379	1.679	1.813
Sell wine by case	Mean	1.00	3.83	3.64	3.10
	Std. Dev.	.000	1.642	1.629	1.826
Sell wine by barrel	Mean	1.00	1.00	2.64	1.60
	Std. Dev.	.000	.000	1.206	1.070
Taste wine for flaws	Mean	2.86	3.75	4.91	3.97
	Std. Dev.	1.676	1.422	.302	1.426
Put together wine list	Mean	2.14	3.33	3.27	3.03
	Std. Dev.	.900	1.614	1.794	1.586
Tour wineries	Mean	3.57	3.50	4.18	3.77
	Std. Dev.	.976	1.000	.874	.971
Help others choose wines to buy	Mean	3.43	4.50	3.91	4.03
	Std. Dev.	.787	.798	1.044	.964
Encourage others to buy particular wine	Mean	2.86	4.33	3.73	3.77
	Std. Dev.	1.215	.888	1.348	1.251
Describe wine making techniques to others	Mean	2.43	3.08	4.09	3.30
	Std. Dev.	1.618	1.443	.539	1.368
Think about disease prevention for grapes	Mean	1.29	1.42	4.18	2.40
	Std. Dev.	.488	.669	1.471	1.694
Attend wine tastings	Mean	4.00	4.42	4.27	4.27
	Std. Dev.	1.155	.900	.786	.907

Conduct wine tastings	Mean	3.57	3.58	3.73	3.63
	Std. Dev.	1.512	1.379	1.272	1.326
Think about pest management	Mean	1.29	1.17	4.27	2.33
	Std. Dev.	.488	.577	1.421	1.768
Sample wines before they are mature	Mean	2.71	2.92	4.64	3.50
	Std. Dev.	1.254	1.084	1.206	1.432
Use heavy machinery	Mean	1.00	1.00	4.00	2.10
	Std. Dev.	.000	.000	1.265	1.647
Research new wine making procedures	Mean	1.43	1.67	3.91	2.43
	Std. Dev.	.535	.888	1.375	1.524
Manage inventory	Mean	1.00	4.17	3.91	3.33
	Std. Dev.	.000	1.403	1.514	1.807
Plan special wine promotion events	Mean	1.57	3.50	3.27	2.97
	Std. Dev.	1.512	1.382	1.489	1.608
Conduct a sensory evaluation of a wine	Mean	3.57	3.17	4.09	3.60
	Std. Dev.	1.618	1.467	1.221	1.429
Conduct a chemical evaluation of a wine	Mean	1.29	1.33	4.18	2.37
	Std. Dev.	.756	.651	1.601	1.771

APPENDIX E:

BACKGROUND QUESTIONNAIRE ACTIVITY INVENTORY

PRINCIPAL COMPONENTS ANALYSIS

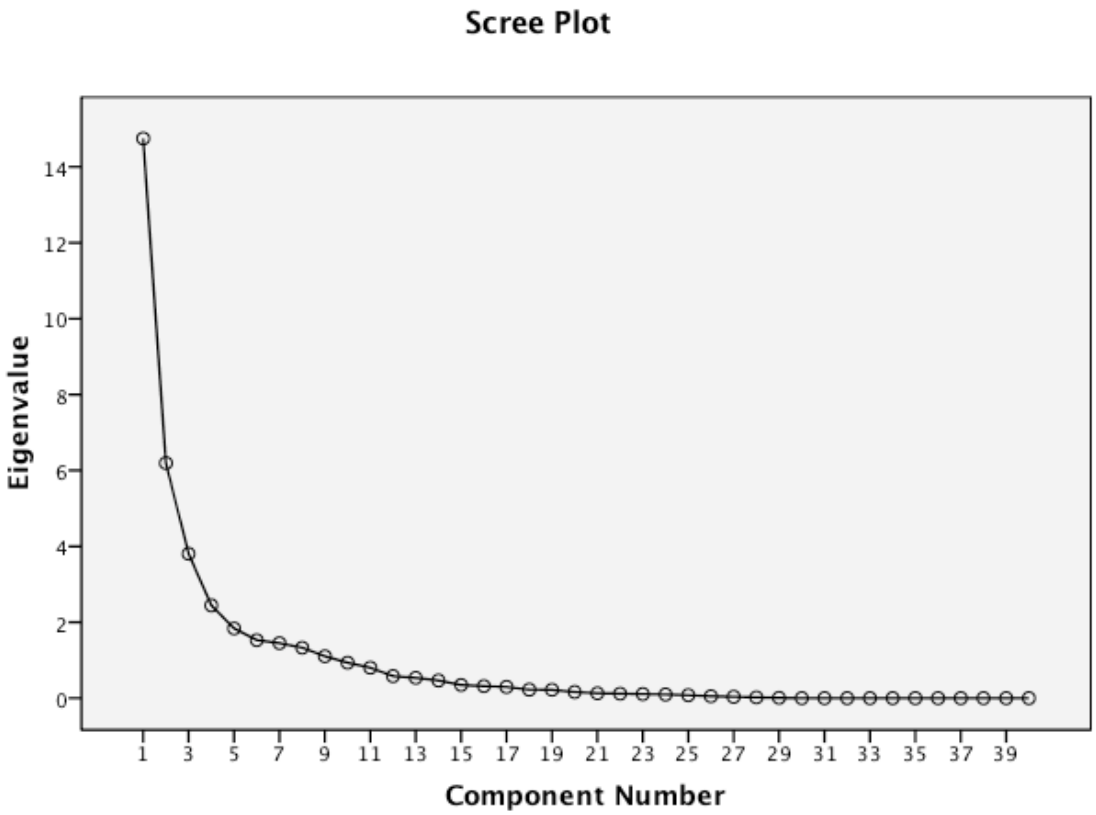


Figure E1 . Scree plot

Table E1. Factor loading scores

Component Matrix

Item	Component				
	1	2	3	4	5
Score a wine using a structured evaluation	.306	.111	-.596	.148	-.017
Give a verbal description of a wine	.626	.539	-.223	.042	-.077
Taste a wine "blind"	.567	.266	-.233	-.313	.006
Write about wine for general public	.538	.259	.016	-.394	.529
Write about wine for specialized audience	.417	.347	-.152	-.372	.123
Speak about wine to groups (10+)	.584	.535	-.112	-.420	-.001
Speak about wine to novices	.618	.604	-.083	-.150	-.152
Speak about wine to knowledgeable audience	.578	.574	-.101	-.066	-.043
Plant or tend grapevines	.696	-.284	-.257	.223	-.247
Harvest grapes	.707	-.255	-.177	.174	-.245
Taste grapes	.851	-.321	-.202	.191	.031
Study wine sales patterns	.704	.066	-.085	.228	-.332
Study wine rankings/ratings	-.126	.559	-.006	.495	.320
Serve wine to customers	.757	-.103	.360	-.181	.256
Set wine prices	.536	-.048	.652	-.057	-.029
Think about wine-food pairings for self or family	.237	.191	-.131	.714	.070

Recommend wine-food pairings to strangers	.538	.614	.156	.133	-.259
Decide when to harvest grapes	.803	-.519	-.159	-.035	-.030
Decide what barrel type to use	.795	-.515	-.150	-.091	-.016
Sell wine by glass	.520	.082	.418	-.204	.322
Sell wine by bottle	.544	.139	.709	.081	-.174
Sell wine by case	.581	.171	.553	.132	-.375
Sell wine by barrel	.663	-.441	.046	-.152	.030
Taste wine for flaws	.658	.070	-.053	.276	.173
Put together wine list	.354	.520	.126	-.168	.092
Tour wineries	.474	.266	-.294	.411	.347
Help others choose wines to buy	.408	.588	.296	.172	.108
Encourage others to buy particular wine	.470	.356	.377	.424	.098
Describe wine making techniques to others	.693	.265	-.351	-.002	-.203
Think about disease prevention for grapes	.855	-.350	-.139	.001	-.048
Attend wine tastings	.318	.666	-.306	-.121	-.124
Conduct wine tastings	.366	.605	-.516	-.216	-.087
Think about pest management	.844	-.384	-.175	-.027	-.004
Sample wines before they are mature	.773	-.344	.032	.335	.157
Use heavy machinery	.800	-.538	.012	-.108	-.050
Research new wine making procedures	.798	-.292	-.064	-.022	.090
Manage inventory	.499	.026	.741	-.016	-.173

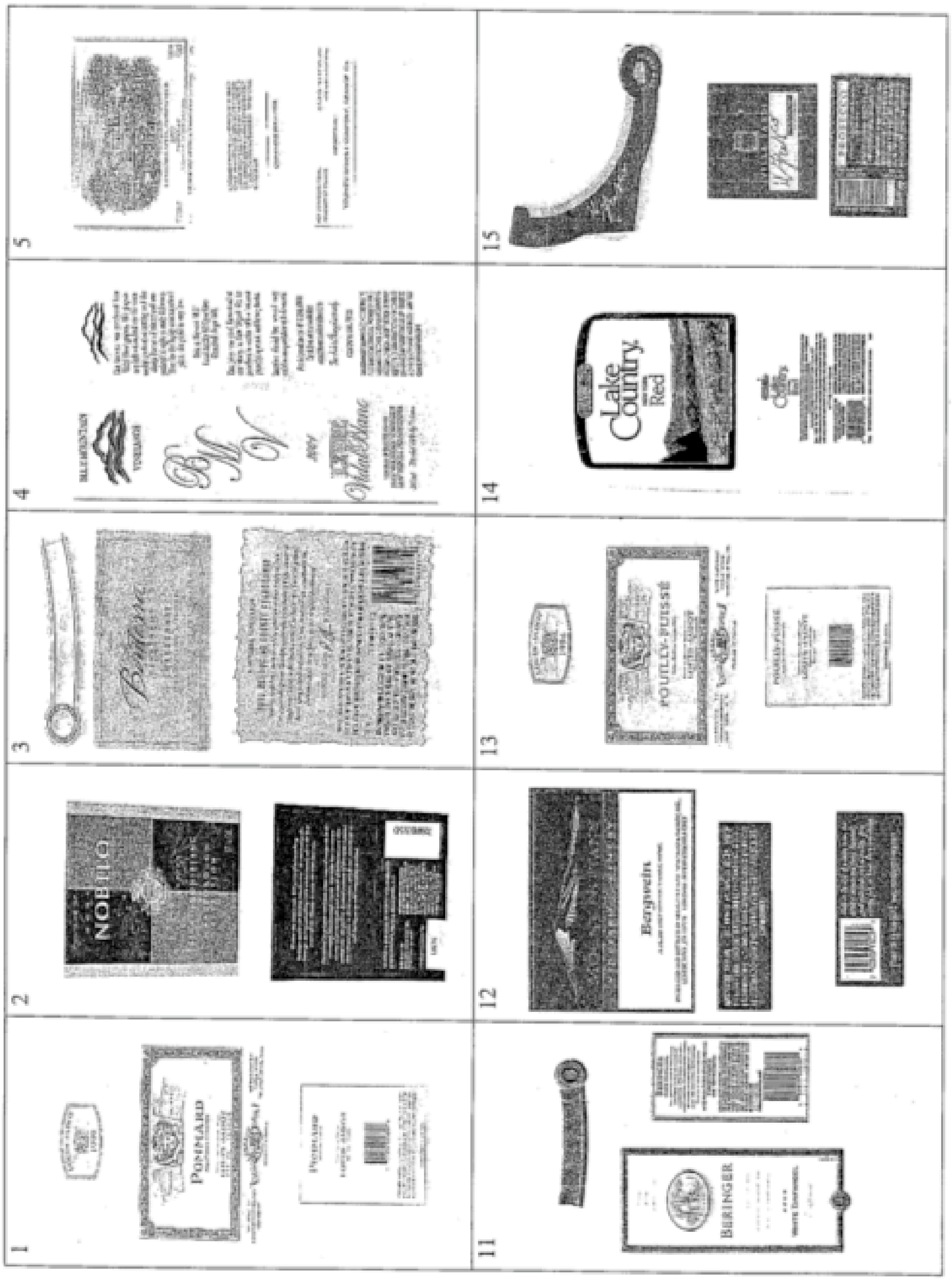
Plan special wine promotion events	.516	.231	.303	-.118	.113
Conduct a sensory evaluation of a wine	.450	-.235	-.050	.117	.661
Conduct a chemical evaluation of a wine	.676	-.552	-.035	-.239	-.052

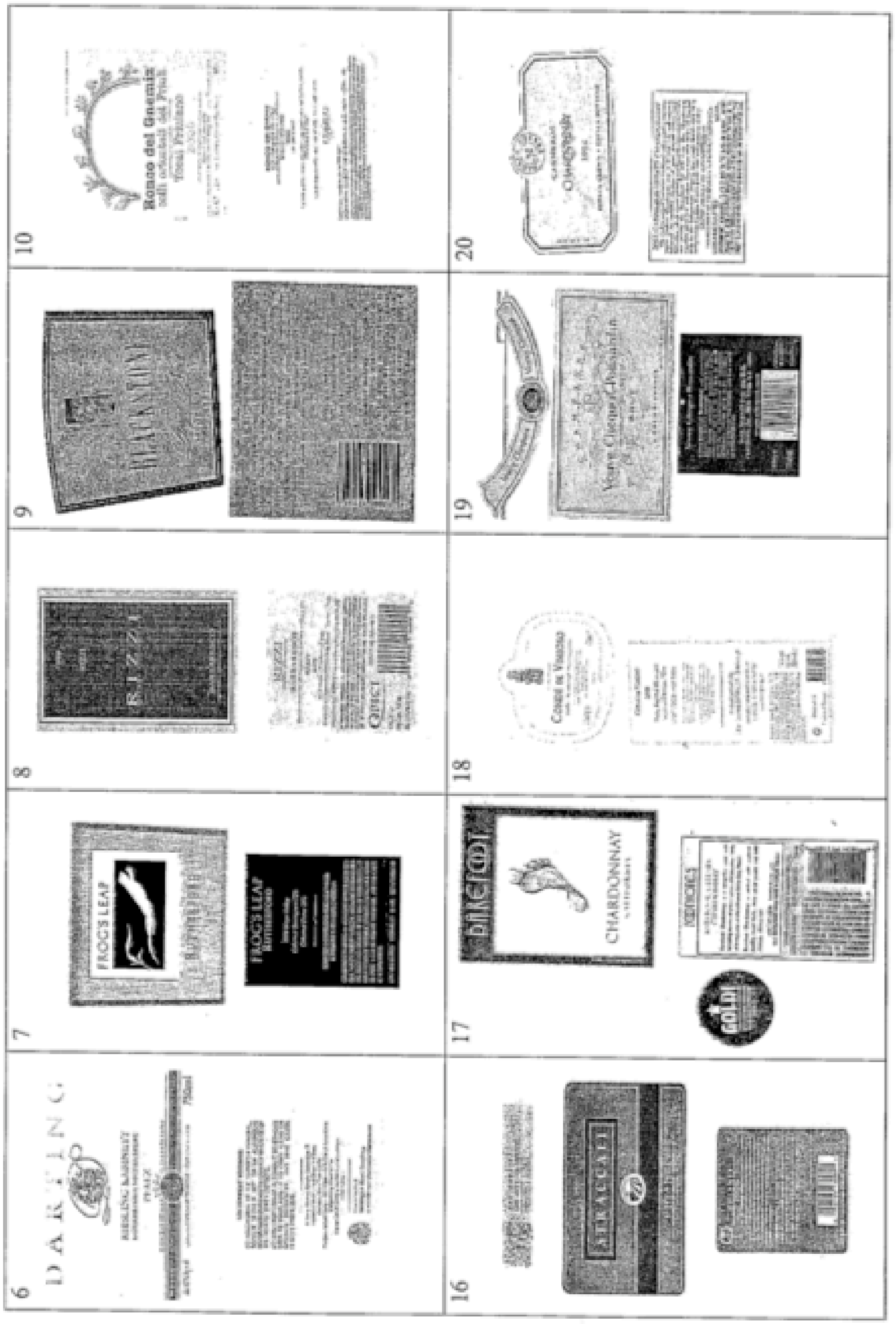
Extraction Method: Principal Component Analysis.

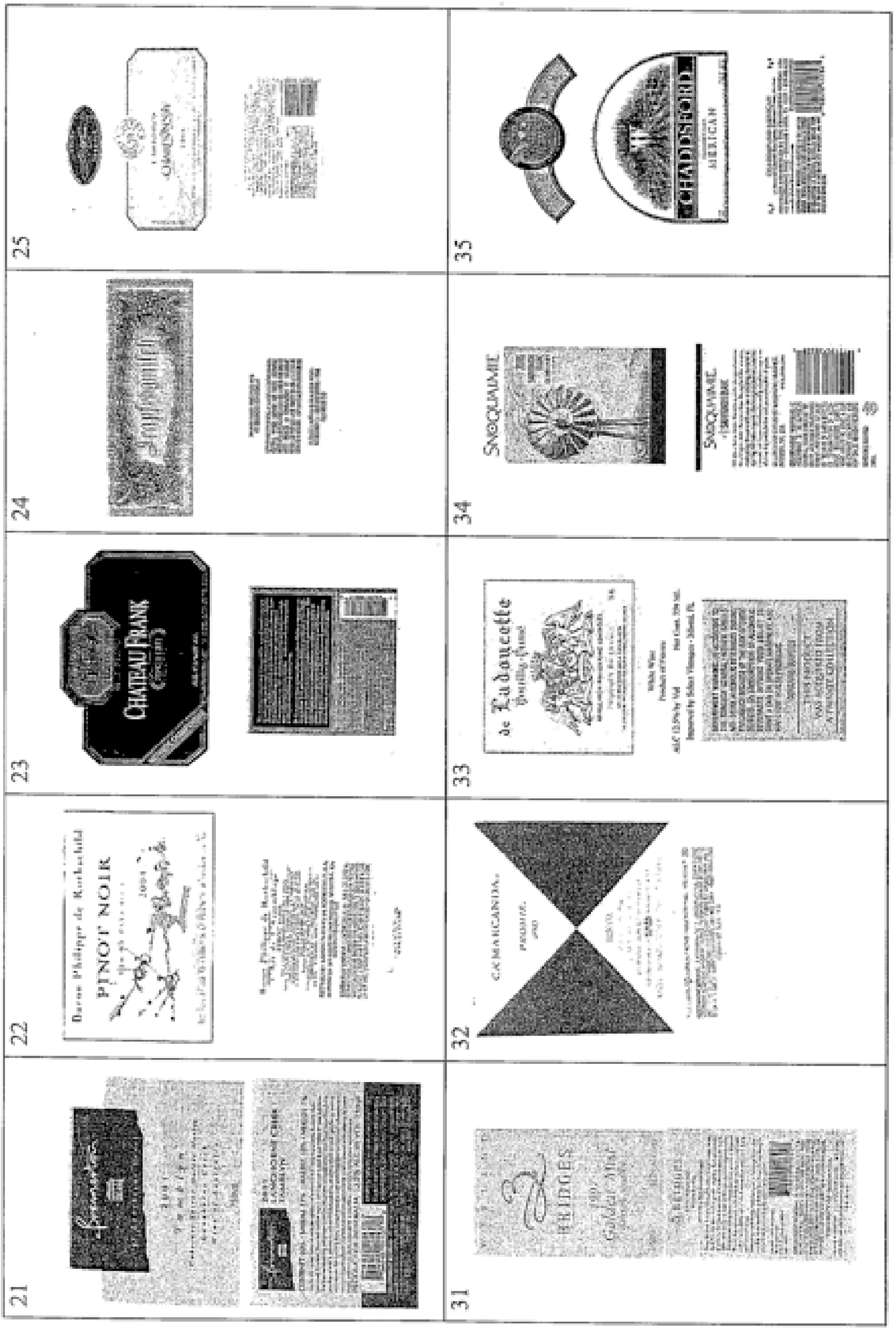
a. 9 components extracted.

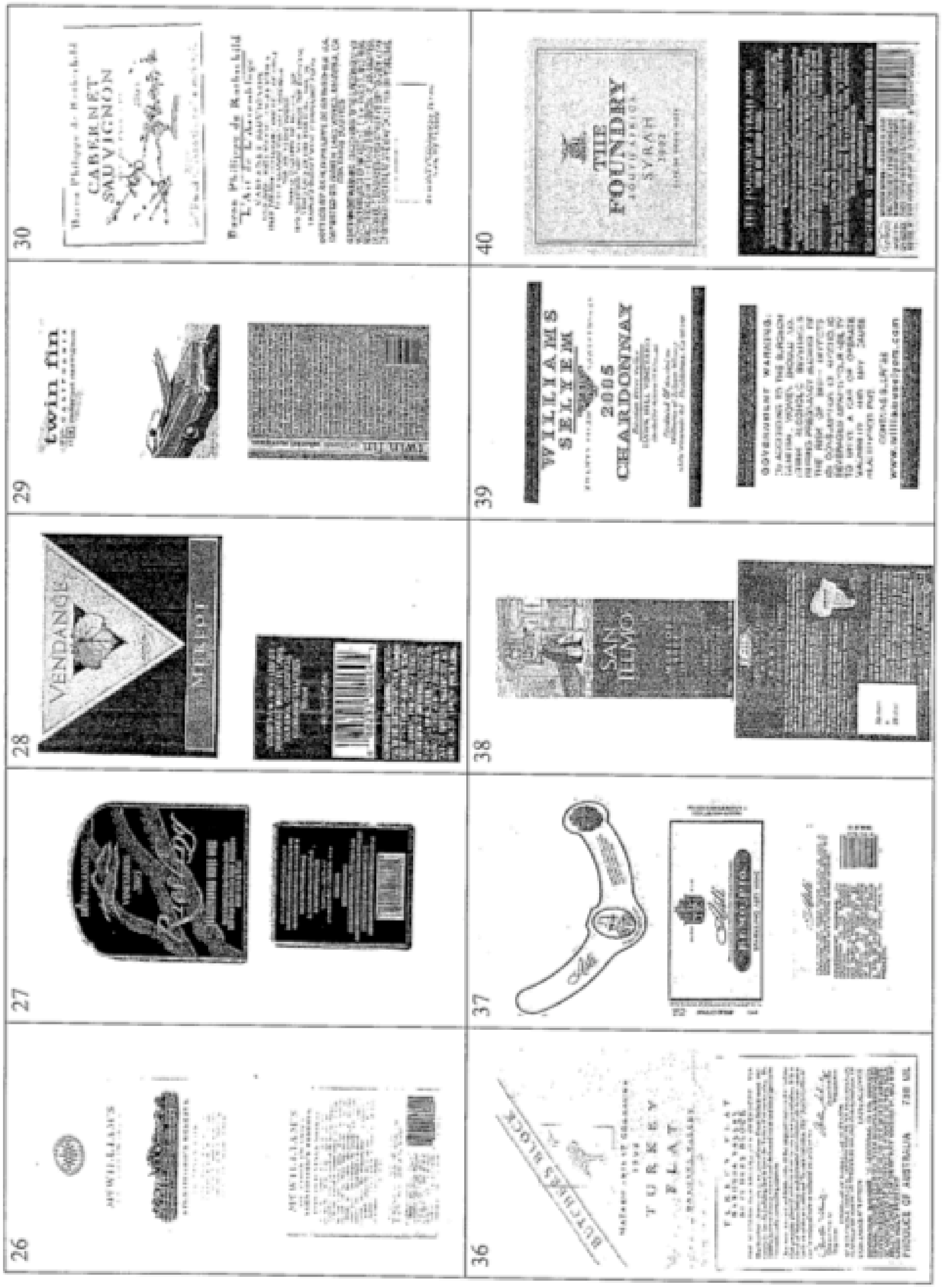
APPENDIX F:

WINE LABEL THUMBNAIL IMAGES






21
22
PINOT NOIR
23
CHATEAU FRANK
24
25
31
32
33
34
SNOQUALMIE
35
CHADDSFORD


26
27
28
29
30
36
37
38
39
40
VENDANGE
MERLOT
twin fin
CABERNET SAUVIGNON
BUTCHERS BLOCK
TURKEY FLAT
WILLIAMS SELYEM
2005
CHARDONNAY
THE FOUNDRY
SYRAH

APPENDIX G:

QUOTA GRIDS

Table 21. Quota grids for wine sampling

	Region				
Type	Old World, French	Old World, Not French	New World, Not US	New World, Western US	New World, Eastern US
Red Table					
White Table					
Other					
Any					
Price	Old World, French	Old World, Not French	New World, Not US	New World, Western US	New World, Eastern US
Low (< $15)					
Medium ($15-$29)					
High (>$29)					

APPENDIX H:

WINE FAMILIARITY BY ITEM AND GROUP

Tables H1, H2, and H3 show the wines that received “familiar” ratings by more than 75% of all experts. The “Hi” column indicates the groups in which at least ¾ of experts said they knew the wine; the “Lo” column indicates the groups in which at least ¾ of experts did not know the wine. The second table contains wines that were familiar to 25-75% of experts overall and the third table contains wines that were familiar to fewer than 25% of experts

Table H1. Mostly familiar wines

#	Wine	C	R	W	All	Hi	Lo
1	Louis Jadot Pommard	90%	90%	100%	93.33%	CRW	
5	Chateau Lafite Rothschild	100%	100%	90%	96.67%	CRW	
7	Frog’s Leap Rutherford	100%	90%	100%	96.67%	CRW	
11	Beringer White Zinfandel	100%	100%	60%	86.67%	CR	
13	Louis Jadot Pouilly Fuisse	100%	100%	90%	96.67%	CRW	
14	Taylor Lake Country (New York) Red	100%	90%	100%	96.67%	CRW	
17	Barefoot Chardonnay	80%	90%	60%	76.67%	CR	
19	Veuve Clicquot Ponsardin Brut	90%	100%	70%	86.67%	CR	
20	Gloria Ferrer Chardonnay	80%	90%	80%	83.33%	CRW	
28	Vendange Merlot	70%	100%	70%	80.00%	R	
39	Williams Selyem Chardonnay (Hawk Hill)	90%	70%	70%	76.67%	C	
35	Chaddsford Merican	100%	80%	100%	93.33%	CRW	

Table H2. Wines of intermediate familiarity

	Wine	C	R	W	All	Hi	Lo
2	Nobilo “Icon” (Marlborough) Sauvignon Blanc	60%	100%	60%	73.33%	R	
3	Bonterra Chardonnay (Mendocino)	50%	100%	40%	63.33%	R	
4	Blue Mountain Icewine Vidal Blanc	30%	70%	70%	56.67%		
6	Darting Riesling Kabinett Durkheimer Michelsberg (Pfalz)	50%	70%	50%	56.67%		
9	Blackstone Pinot Noir (Monterey)	70%	100%	50%	73.33%	R	
12	Mount Nittany Bergwein	20%	50%	90%	53.33%	W	C
16	Straccali Chianti	40%	70%	10%	40.00%		W
22	Baron Philippe de Rothschild Pinot Noir	40%	70%	30%	46.67%		
23	Chateau Frank Brut	30%	70%	80%	60.00%	R	
24	Leapfrogmilch	30%	60%	50%	46.67%		
25	Gloria Ferrer (Carneros) Chardonnay Reserve	80%	70%	70%	73.33%	C	
26	McWilliam’s of Coonawarra (Stentiford’s Reserve) Old Vines Shiraz	30%	60%	10%	33.33%		W
27	Blue Mountain Riesling	30%	60%	80%	56.67%	W	
29	Twin Fin Cabernet Sauvignon	10%	60%	30%	33.33%		W

30	Baron Philippe de Rothschild Cabernet Sauvignon	70%	70%	70%	70.00%		
32	Gaja Ca'Marcanda Promis	10%	70%	10%	30.00%		CW
33	de Ladoucette Pouilly Fume	50%	100%	30%	60.00%	R	
34	Snoqualmie Sauvignon Blanc (Columbia)	60%	60%	40%	53.33%		
36	Turkey Flat Butchers Block Mataro Shiraz Grenache (Barossa)	40%	90%	30%	53.33%	R	
37	Elmo Pio Asti	90%	60%	60%	70.00%	C	

Table H3. Mostly unfamiliar wines

	Mostly Unfamiliar Wines						
	Wine	C	R	W	All	Hi	Lo
8	Rizzi (Azienda Vitivinicola) Barbaresco	30%	30%	10%	23.33%		W
10	Ronco del Gnemiz colli orientali del Friuli Tocai Friulano	10%	40%	0%	16.67%		CW
15	Villa Sandi Prosecco	10%	50%	10%	23.33%		CW
18	Conde de Vimioso	10%	10%	0%	6.67%		
21	Bremerton Tamblyn (Langhorne Creek)	0%	60%	0%	20.00%		CW
31	3 Bridges Golden Mist Botrytis Semillon	10%	10%	0%	6.67%		CRW
38	San Telmo (Mendoza) Merlot	20%	30%	0%	16.67%		

40	The Foundry (Cape of Good Hope) Syrah	10%	40%	10%	20.00%		CW

Note: These tables reflect the initial expert group assignments (10-10-10).

APPENDIX I:

PROCEDURAL COMMENT ON CARD REVIEW TASK

The familiarity and tasting results cannot be taken too literally because participants were inconsistent in the way that they interpreted the questions, “Are you familiar with this wine?” and “Have you ever tasted a bottle of this wine?” Sometimes subjects responded that they had never tasted a wine because they had never tried that particular vintage (although they had tried that same blend from that winemaker and vineyard). On the other hand, sometimes subjects responded that they *had* tasted a wine, despite the fact that they had never heard of the winemaker (they had, however, tried that particular *type* of wine, say, an ice wine). Occasionally, despite responding that they were unfamiliar with the wine, they still asserted that they had tasted it. These differences in interpretation sometimes occurred within the same individual. Had the precise interpretation always been made explicit, it might have been possible to code them consistently, but more often people merely responded “yes” or “no” without elaboration. Although this problem appeared early in data collection, there was no straightforward solution. Most options would have been extremely time-consuming for the subject and probably would have biased them to attend to specific dimensions, which was important to avoid. Therefore, the data are just a rough indication of experts’ experience with the wines.

APPENDIX J:

WINE SET CHARACTERISTICS

The table presents the 40 wines in the set, sorted by color, region, and grape. The prices reported are those set by the PLCB price for spring 2007.

#	Name	Color	Region	Grape varietal(s)	Price
11	Beringer White Zinfandel	Other	US-CA	zinfandel	15.99
38	San Telmo Merlot	Red	Argentina	merlot	$8.99
21	Bremerton Tamblyn	Red	Australia	cabernet sauvignon, shiraz, malbec, merlot	$17.29
26	McWilliam's of Coonawarra Old Vines Shiraz	Red	Australia	shiraz	$15.99
36	Turkey Flat Butchers Block	Red	Australia	shiraz, mataro & grenache	$17.99
30	Baron Philippe de Rothschild Cabernet Sauvignon	Red	France	cabernet sauvignon	$9.99
5	Chateau Lafite Rothschild	Red	France	cabernet sauvignon, merlot, cabernet franc, petit verdot	$299.99
1	Louis Jadot Pommard	Red	France	pinot noir	$29.99
22	Baron Philippe de Rothschild Pinot Noir	Red	France	pinot noir	$5.99
32	Gaja Ca'Marcanda Promis	Red	Italy	merlot, syrah, sangiovese	$33.99
8	Rizzi Barbaresco	Red	Italy	nebbiolo	$24.99
16	Straccali Chianti	Red	Italy	sangiovese, canaiolo, merlot	$8.99

18	Conde de Vimioso	Red	Portugal	blend	$11.69
40	The Foundry Syrah	Red	South Africa	shiraz	$39.99
29	Twin Fin Cabernet Sauvignon	Red	U.S.-CA	cabernet sauvignon	$9.99
7	Frog's Leap Rutherford	Red	U.S.-CA	cabernet sauvignon, cabernet franc	$59.99
28	Vendange Merlot	Red	U.S.-CA	merlot	$4.79
9	Blackstone Pinot Noir	Red	U.S.-CA	pinot noir	$11.99
14	Taylor Lake Country Red	Red	U.S.-NY	blend	$4.99
35	Chaddsford Merican	Red	US-PA	cabernet sauvignon, cabernet franc, merlot	37.99
31	3 Bridges Golden Mist Botrytis	White	Australia	semillon	$12.99
13	Louis Jadot Pouilly Fuisse	White	France	chardonnay	$19.99
19	Veuve Clicquot Ponsardin Brut	White*	France	pinot noir, chardonnay	$43.99
33	de Ladoucette Pouilly Fume	White	France	sauvignon blanc	$12.99
6	Darting Riesling Kabinett Durkheimer Michelsberg	White	Germany	riesling	$13.49
37	Elmo Pio Asti	White*	Italy	moscato	$6.29
15	Villa Sandi Prosecco	White*	Italy	prosecco	$12.99
10	Ronco del Gnemiz Friuli Tocai Friulano	White	Italy	tokai	$12.99
2	Nobilo "Icon" Sauvignon Blanc	White	New Zealand	sauvignon blanc	$17.99
3	Bonterra Chardonnay	White	U.S.-CA	chardonnay	$12.99

17	Barefoot Chardonnay	White	U.S.-CA	chardonnay	$6.99
20	Gloria Ferrer Chardonnay (Estate)	White	U.S.-CA	chardonnay	$10.99
25	Gloria Ferrer Chardonnay (Reserve)	White	U.S.-CA	chardonnay	$16.99
39	Williams Selyem Chardonnay	White	U.S.-CA	chardonnay	$68.99
24	Leapfrogmilch	White	U.S.-CA	riesling, chardonnay	$12.99
23	Chateau Frank Brut	White*	U.S.-NY	pinot noir, chardonnay	$24.99
12	Mount Nittany Bergwein	White	U.S.-PA	blend	$7.99
27	Blue Mountain Riesling	White	U.S.-PA	riesling	$12.69
4	Blue Mountain Icewine	White	U.S.-PA	vidal blanc	$34.99
34	Snoqualmie Sauvignon Blanc	White	U.S.-WA	sauvignon blanc	$9.99

- Sparkling

APPENDIX K:

MULTIDIMENSIONAL SCALING OF WINES FOR RETAILERS AND WINEMAKERS

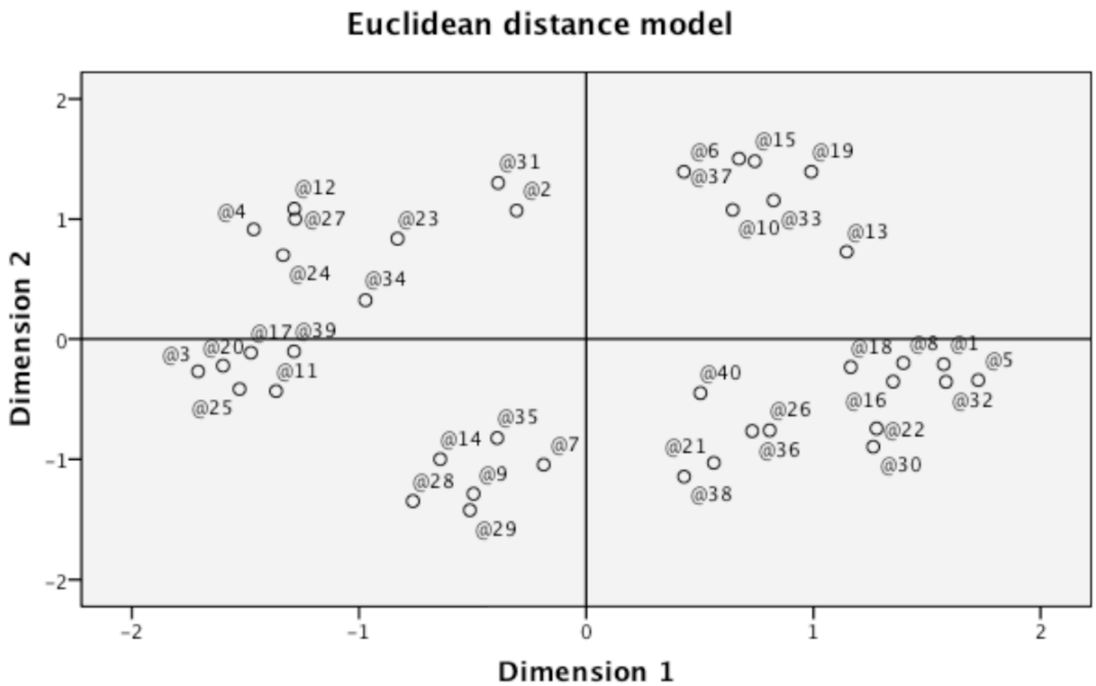


Figure K1. Multidimensional scaling of wines: Retailers

Note: The numbers identify the wines. See Appendix I2 for a list of the wines along with key characteristics. To obtain this plot, the SPSS ALSCAL procedure mapped wines to approximate the distances provided by the averaged distance matrix for all retailers' sorts.

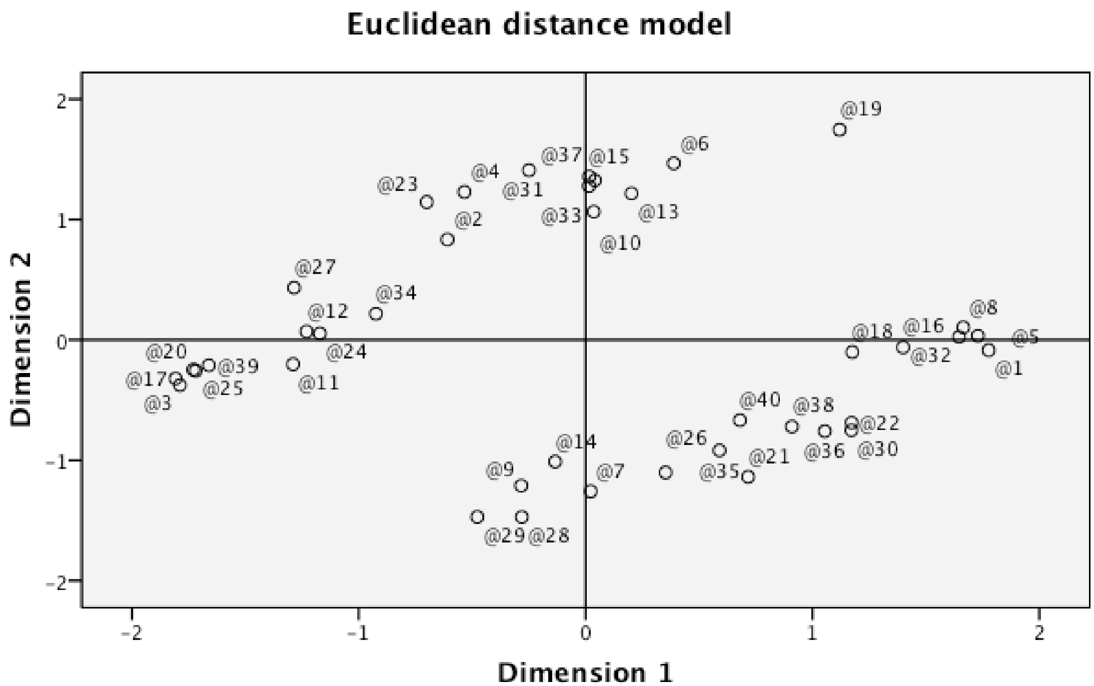


Figure K2. Multidimensional scaling of wines: Winemakers

Note: The numbers identify the wines. See Appendix I2 for a list of the wines along with key characteristics. To obtain this plot, the SPSS ALSCAL procedure mapped wines to approximate the distances provided by the averaged distance matrix for all winemakers' sorts.

APPENDIX L:

THREE-DIMENSIONAL MDS SOLUTION

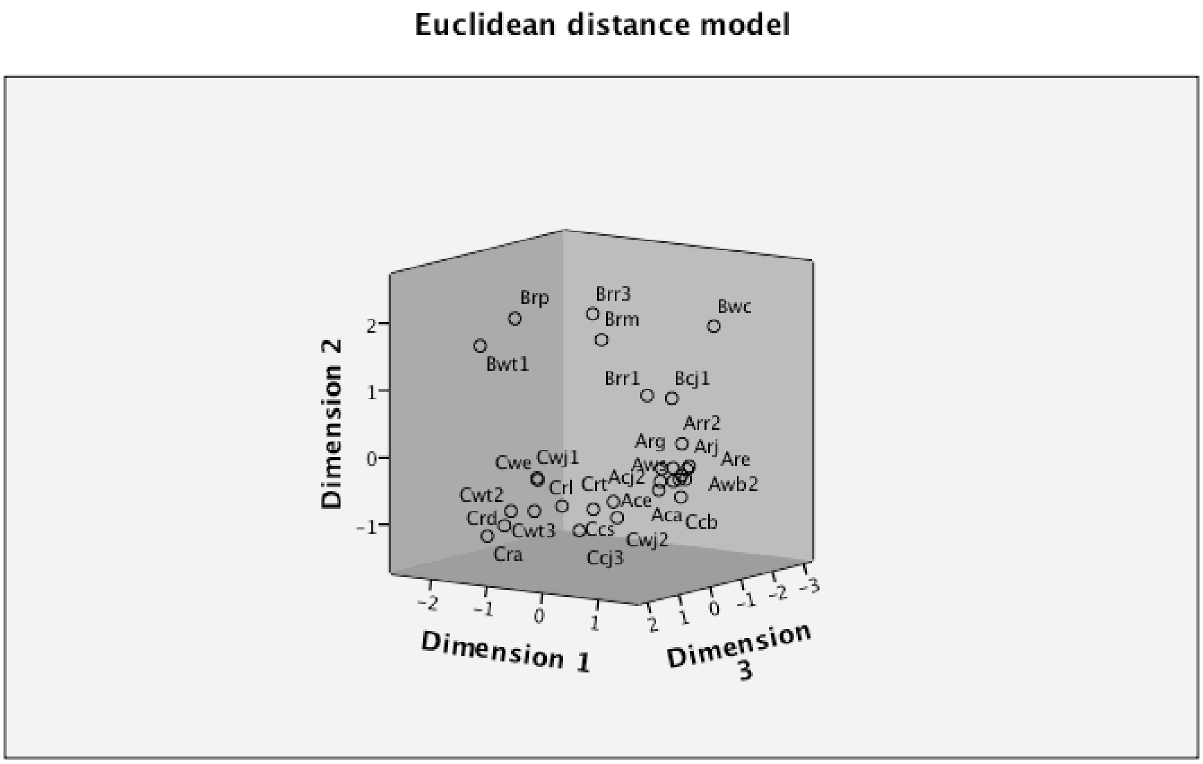


Figure L1. Three-dimensional scaling of standardized distances from experts' first sorts

Note: The first letter indicates group based on the hierarchical cluster analysis presented in Figure 5. The second letter indicates the expert group (connoisseur, retailer, or winemakers). To obtain this plot, the SPSS ALSCAL program calculated Euclidean distances among experts based on the 780 pair-wise distances from each expert's first sort. The distances were standardized using z-scores. RSQ = .90.

APPENDIX M:

PROPORTION OF EXPERTS FROM EACH GROUP WHO USED CATEGORY TYPES

Table M1. Proportion of experts from each group who used given category types

Category type	Connoisseurs *n* = 7	Retailers *n* = 12	Winemakers *n* = 11	All experts
grape	1.00	0.75	0.73	0.80
color	1.00	0.92	0.91	0.93
preference	0.14	0.25	0.64	0.37
price/quality	1.00	0.92	0.73	0.87
process	0.86	0.67	0.91	0.80
region	1.00	1.00	1.00	1.00
role	0.57	0.75	0.36	0.57
style	1.00	0.92	0.91	0.93
type	1.00	0.92	0.82	0.90

APPENDIX N:

ACKNOWLEDGEMENT FORM

Northwestern University Psychology Department

Project Title:
Domain knowledge and reasoning among wine experts

Student Investigator	**Faculty Advisor**
Julia Beth Proffitt	**Douglas L. Medin**

Your assistance has been extremely valuable to me and I would like to recognize your contribution in a way that is acceptable to you. When I publish the results of these studies, I intend to thank the participants in the acknowledgments. Which best describes your preference:

- ❒ Please do not identify me personally.
  If you choose this, you will be thanked only in a general way, e.g., "three winemakers".

- ❒ You may identify me by name in the acknowledgments.
  If you choose this, your name will be listed in the acknowledgments, but nowhere else. Your responses and comments will remain anonymous.

- ❒ Other: ____

  ____
  If your preference is different, please describe it in the space above and I will do my best to respect it.

____________________________	____________________________	____________
Name (printed)	Signature	Date

If you would like to receive a copy of the published results, please print your mailing address here:

____

____

APPENDIX O:

PARTICIPANT DEBRIEFING

Thank you for participating in this study. Your expertise and time are greatly appreciated.

During this research, you were asked to describe your wine-related background and share your knowledge about the domain by grouping wines into categories and evaluating statements about wine. The primary purpose of this research is to develop a better understanding of how people organize and use knowledge. Research on categories and concepts is central to that goal, but most studies of categories have focused on the learning of new, artificial categories.

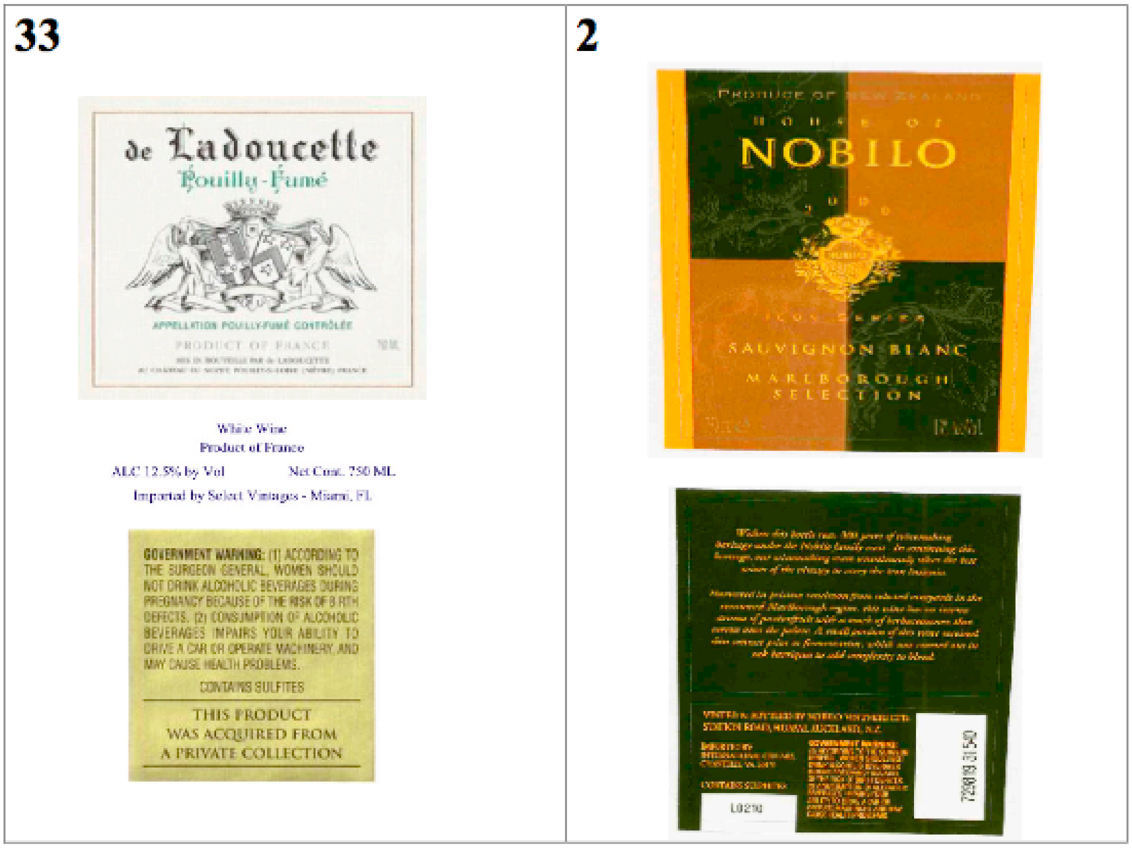
We are interested in the structure of categories that have been acquired naturally, through extensive experience in a domain. Furthermore, we expect that different kinds of experts may tend to organize the domain in different ways, based on differences in how they typically interact with it. Although sommeliers and winemakers might both know a lot about a given wine, we predict that they may tend to highlight different features and see it as a member of different "families". In addition, this research furthers work exploring how concepts are used to make inferences.

Finally, we hope to contribute to our understanding of wine expertise, which has focused on experts' ability to perceive, describe, and remember the sensory features of a wine. Little systematic attention has been paid to how wine experts organize and use their substantial knowledge base about winemaking and viticulture. We believe that this "contextual" knowledge is, in fact, central.

Thank you again for your participation in this research. If you have any questions, your experimenter will be happy to answer them. Or, if you think of any questions or comments later, please contact her (Beth Proffitt, juliabeth.proffitt@fandm.edu, 717-291-3950) or the principal investigator (Dr. Douglas Medin, medin@northwestern.edu, 847-467-1150).

APPENDIX P:

TWO SAMPLE SIMILARITY TRIALS

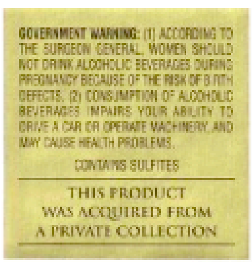


How similar are these two wines?

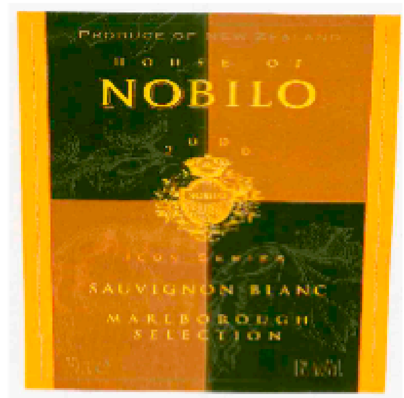
Sauvignon blanc wines

33

White Wine
Product of France
ALC 12.5% by Vol Net Cont. 750 ML
Imported by Select Vintages - Miami, FL



2



How similar are these two sauvignon blanc wines?

APPENDIX Q:

SAMPLE SIMILARITY RATING SHEET

Similarity Rating Sheet

Sheet	**Pair**	**Not at All Similar**						**Very Similar**
1	(2, 33)	1	2	3	4	5	6	7
2	(17, 39)	1	2	3	4	5	6	7
3	(9, 13)	1	2	3	4	5	6	7
4	(7, 19)	1	2	3	4	5	6	7
5	(8, 16)	1	2	3	4	5	6	7
6	(5, 35)	1	2	3	4	5	6	7
7	(5, 6)	1	2	3	4	5	6	7
8	(2, 12)	1	2	3	4	5	6	7
9	(1, 28)	1	2	3	4	5	6	7
10	(7, 20)	1	2	3	4	5	6	7
11	(6, 11)	1	2	3	4	5	6	7
12	(23, 37)	1	2	3	4	5	6	7
13	(11, 29)	1	2	3	4	5	6	7

APPENDIX R:

CATEGORIES USED IN CATEGORY MEMBERSHIP TASK

Block 1	Wines
Block 2	Bordeaux Blend Wines Sauvignon Blanc Wines Sparkling Wines
Block 3	Italian Wines Light White Wines Party Wines
Block 4	Old World Wines Single Varietal Wines
Block 5	High-End Wines Champagne Wines Sweet Wines

APPENDIX S:

CATEGORY MEMBERSHIP AGREEMENT

	Bordeaux Blend				Champagne			
	C	R	W	Total	C	R	W	Total
1 Louis Jadot Pommard	0.60	0.60	0.56	0.58	1	1	1	1
2 Nobilo "Icon" Sauvignon Blanc	1	0.60	0.78	0.75	1	0.75	1	0.90
3 Bonterra Chardonnay	1	0.80	0.78	0.83	1	0.75	1	0.90
4 Blue Mountain Icewine	1	0.80	0.78	0.83	1	1	1	1
5 Chateau Lafite Rothschild	1	1	1	1	1	1	1	1
6 Darting Riesling Kabinett	1	1	0.78	0.92	1	1	1	1
7 Frog's Leap Rutherford	0.60	1	0.78	0.83	1	1	1	1
8 Rizzi Barbaresco	1	0.40	0.56	0.58	1	1	1	1
9 Blackstone Pinot Noir	1	0.60	0.56	0.67	1	1	1	1
10 Gnemiz Tocai Friulano	1	1	0.78	0.92	1	1	1	1
11 Beringer White Zinfandel	1	0.80	1	0.92	1	1	1	1
12 Mount Nittany Bergwein	1	1	0.78	0.92	1	1	1	1
13 Louis Jadot Pouilly Fuisse	1	0.80	0.56	0.75	1	0.75	1	0.90
14 Taylor Lake Country Red	1	0.60	0.56	0.67	1	1	1	1
15 Villa Sandi Prosecco	1	0.80	1	0.92	0	0.25	0.11	0.05
16 Straccali Chianti	1	0.40	0.56	0.58	1	1	1	1
17 Barefoot Chardonnay	1	0.80	0.78	0.83	1	0.75	1	0.90
18 Conde de Vimioso	0.60	0.60	0.33	0.50	1	1	1	1
19 Veuve Clicquot Ponsardin Brut	1	0.80	0.78	0.83	1	1	1	1
20 Gloria Ferrer Chard. (Estate)	1	0.80	0.78	0.83	1	0.75	1	0.90
21 Bremerton Tamblyn	0.20	0.60	0.33	0.42	1	1	1	1
22 B.P. Rothschild Pinot Noir	0.60	0.60	0.56	0.58	1	1	1	1
23 Chateau Frank Brut	1	1	0.78	0.92	0.50	0	0.56	0.33
24 Leapfrogmilch	1	0.80	0.78	0.83	1	1	1	1
25 Gloria Ferrer Chard. (Reserve)	1	0.80	0.78	0.83	1	0.75	1	0.90
26 McWilliam's Shiraz	1	0.20	0.11	0.33	1	1	1	1
27 Blue Mountain Riesling	1	0.80	0.78	0.83	1	1	1	1
28 Vendange Merlot	0.60	0.20	0.11	0.25	1	1	1	1
29 Twin Fin Cabernet Sauvignon	0.60	0.20	0.11	0.08	1	1	1	1
30 B.P. Rothschild Cab. Sauv.	0.60	0.40	0.11	0.33	1	1	1	1
31 3 Bridges Botrytis Semillon	1	0.60	0.78	0.75	1	1	1	1
32 Gaja Ca'Marcanda Promis	0.60	0.20	0.56	0.42	1	1	0.78	0.90
33 de Ladoucette Pouilly Fume	1	0.40	0.56	0.58	1	0.75	1	0.90
34 Snoqualmie Sauvignon Blanc	0.60	0.60	0.78	0.67	1	0.75	1	0.90
35 Chaddsford Merican	0.20	0	1	0.33	1	1	1	1
36 Turkey Flat Butchers Block	1	0.20	0.11	0.33	1	1	1	1
37 Elmo Pio Asti	1	1	1	1	0.50	0.25	0.11	0.05
38 San Telmo Merlot	0.60	0.20	0.11	0.08	1	1	1	1
39 Williams Selyem Chardonnay	1	0.80	0.78	0.83	1	0.75	1	0.90
40 The Foundry Syrah	1	0.60	0.78	0.75	1	1	1	1
	0.86	0.64	0.63	0.67	0.95	0.88	0.94	0.91

	High End				Italian			
	C	R	W	Total	C	R	W	Total
1 Louis Jadot Pommard	1	1	1	1	1	0.60	1	0.83
2 Nobilo "Icon" Sauvignon Blanc	0	0.50	0.56	0.43	1	0.60	0.78	0.75
3 Bonterra Chardonnay	0.50	0.50	0.56	0.33	1	0.60	1	0.83
4 Blue Mountain Icewine	1	0.75	0.78	0.43	0.60	0.60	1	0.75
5 Chateau Lafite Rothschild	1	1	1	1	1	0.80	1	0.92
6 Darting Riesling Kabinett	0.50	0.25	0.56	0.43	1	0.80	1	0.92
7 Frog's Leap Rutherford	0.50	1	0.78	0.81	1	0.80	1	0.92
8 Rizzi Barbaresco	0	1	0.78	0.71	0.60	1	1	0.92
9 Blackstone Pinot Noir	0.50	0.50	0.56	0.33	1	0.60	1	0.83
10 Gnemiz Tocai Friulano	0.50	0.75	0.56	0.43	1	1	0.78	0.92
11 Beringer White Zinfandel	1	0.50	0.33	0.52	1	0.80	0.78	0.83
12 Mount Nittany Bergwein	1	0	0.11	0.24	1	0.60	0.78	0.75
13 Louis Jadot Pouilly Fuisse	1	1	0.56	0.81	1	0.80	1	0.92
14 Taylor Lake Country Red	1	0.50	0.33	0.52	1	0.80	0.78	0.83
15 Villa Sandi Prosecco	0.50	0	0.56	0.14	1	1	0.78	0.92
16 Straccali Chianti	0.50	0.25	0.56	0.43	1	1	1	1
17 Barefoot Chardonnay	1	0	0.33	0.05	1	0.60	1	0.83
18 Conde de Vimioso	0.50	0	0.78	0.24	1	0.20	0.33	0.42
19 Veuve Clicquot Ponsardin Brut	1	1	1	1	1	0.60	1	0.83
20 Gloria Ferrer Chard. (Estate)	0.50	0.75	0.56	0.62	1	0.60	1	0.83
21 Bremerton Tamblyn	0	0.50	0.56	0.43	1	0.80	1	0.92
22 B.P. Rothschild Pinot Noir	0.50	0.25	0.56	0.24	1	0.60	1	0.83
23 Chateau Frank Brut	1	0.25	0.56	0.14	1	0.60	1	0.83
24 Leapfrogmilch	0.50	0.50	0.11	0.14	1	0.80	1	0.92
25 Gloria Ferrer Chard. (Reserve)	0.50	0.50	0.78	0.62	1	0.60	1	0.83
26 McWilliam's Shiraz	0.50	1	1	0.90	1	0.80	1	0.92
27 Blue Mountain Riesling	1	0	0.56	0.05	0.60	0.80	1	0.83
28 Vendange Merlot	0.50	0.50	0.33	0.14	1	0.80	1	0.92
29 Twin Fin Cabernet Sauvignon	1	0.25	0.33	0.14	1	0.60	1	0.83
30 B.P. Rothschild Cab. Sauv.	0.50	0.25	0.56	0.24	1	0.80	1	0.92
31 3 Bridges Botrytis Semillon	0	0.50	0.78	0.52	1	0.60	1	0.83
32 Gaja Ca'Marcanda Promis	0	1	1	0.81	1	0.80	1	0.92
33 de Ladoucette Pouilly Fume	1	1	1	1	1	0.60	1	0.83
34 Snoqualmie Sauvignon Blanc	0	0.50	0.56	0.43	1	0.60	1	0.83
35 Chaddsford Merican	1	0	0.78	0.14	1	0.80	0.78	0.83
36 Turkey Flat Butchers Block	0.50	0.50	0.33	0.43	1	0.80	1	0.92
37 Elmo Pio Asti	1	0	0.11	0.14	1	1	1	1
38 San Telmo Merlot	0	0	0.56	0.24	1	0.60	1	0.83
39 Williams Selyem Chardonnay	1	0.75	1	0.90	1	0.60	1	0.83
40 The Foundry Syrah	0.50	0.75	0.78	0.71	1	0.80	1	0.92
	0.61	0.51	0.61	0.47	0.97	0.72	0.94	0.86

	Light White				Old World			
	C	R	W	Total	C	R	W	Total
1 Louis Jadot Pommard	1	0.80	1	0.92	0.60	1	1	0.92
2 Nobilo "Icon" Sauvignon Blanc	0.60	1	0.56	0.75	1	0.60	0.78	0.75
3 Bonterra Chardonnay	0.60	0.40	0.56	0.50	1	0.60	0.56	0.67
4 Blue Mountain Icewine	0.60	0	0.11	0.17	1	0.60	0.56	0.67
5 Chateau Lafite Rothschild	1	0.80	1	0.92	1	1	1	1
6 Darting Riesling Kabinett	1	0.60	1	0.83	0.20	0.80	0.78	0.67
7 Frog's Leap Rutherford	1	0.80	1	0.92	0.60	0.60	0.56	0.58
8 Rizzi Barbaresco	1	1	1	1	0.60	1	1	0.92
9 Blackstone Pinot Noir	1	0.80	0.78	0.83	1	0.60	0.78	0.75
10 Gnemiz Tocai Friulano	0.20	0.80	0.56	0.58	0.60	0.80	1	0.83
11 Beringer White Zinfandel	0.60	0.40	0.33	0.08	1	0.60	1	0.83
12 Mount Nittany Bergwein	0.60	0.80	0.56	0.67	1	0.60	1	0.83
13 Louis Jadot Pouilly Fuisse	0.20	0.40	0.33	0.33	0.60	1	1	0.92
14 Taylor Lake Country Red	1	1	0.78	0.92	1	0.60	1	0.83
15 Villa Sandi Prosecco	0.60	1	0.56	0.75	0.60	1	1	0.92
16 Straccali Chianti	1	1	1	1	0.60	1	1	0.92
17 Barefoot Chardonnay	1	0.60	0.33	0.58	1	0.60	0.78	0.75
18 Conde de Vimioso	1	1	1	1	0.60	1	0.78	0.83
19 Veuve Clicquot Ponsardin Brut	0.60	0.40	0.33	0.42	0.60	1	1	0.92
20 Gloria Ferrer Chard. (Estate)	0.60	0.40	0.33	0.42	1	0.60	0.56	0.67
21 Bremerton Tamblyn	1	1	1	1	1	0.60	0.78	0.75
22 B.P. Rothschild Pinot Noir	1	0.80	1	0.92	0.60	0.80	0.78	0.75
23 Chateau Frank Brut	0.60	0.40	0.56	0.50	1	0.60	0.56	0.67
24 Leapfrogmilch	1	0.80	0.56	0.75	0.60	0.60	0.78	0.67
25 Gloria Ferrer Chard. (Reserve)	0.60	0.40	0.33	0.42	1	0.60	0.56	0.67
26 McWilliam's Shiraz	1	1	1	1	0.60	0.60	0.78	0.67
27 Blue Mountain Riesling	1	0.80	0.78	0.83	1	0.60	0.78	0.75
28 Vendange Merlot	1	0.80	1	0.92	1	0.60	0.78	0.75
29 Twin Fin Cabernet Sauvignon	1	0.80	1	0.92	1	0.60	0.78	0.75
30 B.P. Rothschild Cab. Sauv.	1	1	1	1	0.60	0.80	0.78	0.75
31 3 Bridges Botrytis Semillon	0.60	0	0.33	0.25	1	0.60	0.56	0.67
32 Gaja Ca'Marcanda Promis	1	0.80	1	0.92	0.60	1	0.78	0.83
33 de Ladoucette Pouilly Fume	0.60	0.60	0.56	0.58	0.60	1	1	0.92
34 Snoqualmie Sauvignon Blanc	1	1	0.56	0.83	1	0.60	0.56	0.67
35 Chaddsford Merican	1	1	1	1	1	0.60	0.56	0.67
36 Turkey Flat Butchers Block	1	0.80	1	0.92	1	0.60	0.78	0.75
37 Elmo Pio Asti	0.60	0.80	0.56	0.67	0.60	1	1	0.92
38 San Telmo Merlot	1	1	1	1	0.60	0.60	0.56	0.58
39 Williams Selyem Chardonnay	0.60	0.20	0.33	0.33	1	0.60	0.78	0.75
40 The Foundry Syrah	1	0.80	1	0.92	0.60	0.60	0.78	0.67
	0.82	0.72	0.72	0.73	0.80	0.73	0.79	0.77

		Party				Sauvignon Blanc			
		C	R	W	Total	C	R	W	Total
1	Louis Jadot Pommard	0.60	0.20	0.11	0.08	1	1	1	1
2	Nobilo "Icon" Sauvignon Blanc	0.60	0.60	0.33	0.50	1	1	1	1
3	Bonterra Chardonnay	0.60	0.40	0.11	0.33	1	0.80	0.56	0.75
4	Blue Mountain Icewine	0.20	0.20	0.33	0.08	1	0.60	0.56	0.67
5	Chateau Lafite Rothschild	0.60	0.40	0.11	0	1	1	1	1
6	Darting Riesling Kabinett	0.60	0	0.78	0.42	1	0.80	0.56	0.75
7	Frog's Leap Rutherford	0.60	0	0.11	0.17	1	1	1	1
8	Rizzi Barbaresco	0.60	0.20	0.33	0.17	1	1	1	1
9	Blackstone Pinot Noir	0.60	1	0.11	0.58	1	1	1	1
10	Gnemiz Tocai Friulano	0.60	0.40	0.56	0.50	0.20	0.60	0.56	0.50
11	Beringer White Zinfandel	1	0.80	1	0.92	1	1	0.56	0.83
12	Mount Nittany Bergwein	0.60	0.20	1	0.58	1	0.60	0.56	0.67
13	Louis Jadot Pouilly Fuisse	0.60	0.20	0.33	0.33	0.20	0.60	0.11	0.25
14	Taylor Lake Country Red	0.60	0.20	1	0.58	1	1	1	1
15	Villa Sandi Prosecco	1	0.80	0.78	0.83	0.60	0.60	0.56	0.58
16	Straccali Chianti	0.60	1	0.11	0.58	1	1	1	1
17	Barefoot Chardonnay	1	1	0.56	0.83	1	0.80	0.56	0.75
18	Conde de Vimioso	0.60	0.20	0.11	0.25	1	1	1	1
19	Veuve Clicquot Ponsardin Brut	1	0.80	0.56	0.75	1	0.80	0.56	0.75
20	Gloria Ferrer Chard. (Estate)	0.60	0.20	0.11	0.25	1	0.80	0.56	0.75
21	Bremerton Tamblyn	0.60	0.40	0.11	0.33	1	1	1	1
22	B.P. Rothschild Pinot Noir	0.60	0.80	0.33	0.58	1	1	1	1
23	Chateau Frank Brut	1	0.40	0.56	0.58	1	0.60	0.56	0.67
24	Leapfrogmilch	1	0.40	1	0.75	1	0.60	0.56	0.67
25	Gloria Ferrer Chard. (Reserve)	0.60	0.20	0.11	0.25	1	0.80	0.56	0.75
26	McWilliam's Shiraz	0.60	0.20	0.11	0.25	1	1	1	1
27	Blue Mountain Riesling	0.60	0.20	1	0.58	1	0.60	0.56	0.67
28	Vendange Merlot	0.60	0.80	0.33	0.58	1	1	1	1
29	Twin Fin Cabernet Sauvignon	0.60	1	0.56	0.75	1	1	1	1
30	B.P. Rothschild Cab. Sauv.	0.60	1	0.11	0.58	1	1	1	1
31	3 Bridges Botrytis Semillon	0.20	0.20	0.33	0.08	1	0.60	0.33	0.58
32	Gaja Ca'Marcanda Promis	0.60	0.20	0.11	0.08	1	0.80	1	0.92
33	de Ladoucette Pouilly Fume	0.60	0	0.33	0.25	1	0.60	1	0.83
34	Snoqualmie Sauvignon Blanc	0.60	0.20	0.33	0.33	1	1	1	1
35	Chaddsford Merican	0.60	0.20	0.11	0.08	1	1	1	1
36	Turkey Flat Butchers Block	0.60	0.60	0.11	0.42	1	1	1	1
37	Elmo Pio Asti	1	0.80	0.78	0.83	1	0.60	0.56	0.67
38	San Telmo Merlot	0.60	0.40	0.33	0.42	1	1	1	1
39	Williams Selyem Chardonnay	0.60	0	0.11	0.17	1	0.80	0.56	0.75
40	The Foundry Syrah	0.60	0	0.33	0.25	1	1	1	1
		0.65	0.42	0.39	0.42	0.95	0.85	0.78	0.84

	Single Varietal				Sparkling			
	C	R	W	Total	C	R	W	Total
1 Louis Jadot Pommard	0.60	0.80	1	0.83	1	0.80	1	0.92
2 Nobilo "Icon" Sauvignon Blanc	1	1	1	1	1	0.80	1	0.92
3 Bonterra Chardonnay	1	1	1	1	0.60	0.80	1	0.83
4 Blue Mountain Icewine	1	0.80	1	0.92	1	0.80	0.78	0.83
5 Chateau Lafite Rothschild	0.60	0.20	0.33	0.33	1	1	1	1
6 Darting Riesling Kabinett	1	1	1	1	1	1	0.56	0.83
7 Frog's Leap Rutherford	1	0.40	0.33	0.50	1	0.80	1	0.92
8 Rizzi Barbaresco	0.60	0.80	0.56	0.67	1	0.80	1	0.92
9 Blackstone Pinot Noir	1	0.80	1	0.92	1	0.80	1	0.92
10 Gnemiz Tocai Friulano	0.20	1	0.78	0.67	1	0.80	0.78	0.83
11 Beringer White Zinfandel	1	0.40	0.78	0.67	1	1	0.78	0.92
12 Mount Nittany Bergwein	0.60	0.20	0.56	0.42	1	0.80	0.56	0.75
13 Louis Jadot Pouilly Fuisse	1	1	1	1	0.60	0.80	0.78	0.75
14 Taylor Lake Country Red	0.60	0.20	1	0.58	1	1	1	1
15 Villa Sandi Prosecco	0.20	0.80	0.11	0.42	1	1	1	1
16 Straccali Chianti	0.20	0.20	0.33	0	1	1	1	1
17 Barefoot Chardonnay	1	1	1	1	1	0.80	1	0.92
18 Conde de Vimioso	0.20	0.20	0.11	0	1	0.80	1	0.92
19 Veuve Clicquot Ponsardin Brut	0.20	0.40	0.11	0.25	1	1	1	1
20 Gloria Ferrer Chard. (Estate)	1	1	1	1	0.60	0.80	1	0.83
21 Bremerton Tamblyn	1	0.40	0.33	0.50	1	1	1	1
22 B.P. Rothschild Pinot Noir	1	0.80	1	0.92	0.60	0.80	1	0.83
23 Chateau Frank Brut	1	0.40	0.33	0.50	1	1	1	1
24 Leapfrogmilch	1	0.40	0.78	0.67	1	1	0.56	0.83
25 Gloria Ferrer Chard. (Reserve)	1	1	1	1	0.60	0.80	1	0.83
26 McWilliam's Shiraz	1	0.80	1	0.92	1	0.80	1	0.92
27 Blue Mountain Riesling	1	1	1	1	1	1	0.56	0.83
28 Vendange Merlot	1	0.60	1	0.83	1	1	1	1
29 Twin Fin Cabernet Sauvignon	0.60	0.80	1	0.83	1	0.80	1	0.92
30 B.P. Rothschild Cab. Sauv.	0.60	0.80	1	0.83	1	0.80	1	0.92
31 3 Bridges Botrytis Semillon	1	1	1	1	1	0.80	0.78	0.83
32 Gaja Ca'Marcanda Promis	0.20	0	0.11	0.08	1	0.80	1	0.92
33 de Ladoucette Pouilly Fume	0.20	1	0.78	0.75	1	0.80	0.78	0.83
34 Snoqualmie Sauvignon Blanc	1	1	1	1	1	0.80	1	0.92
35 Chaddsford Merican	1	0.40	0.33	0.50	1	1	1	1
36 Turkey Flat Butchers Block	1	0.20	0.33	0.42	1	1	1	1
37 Elmo Pio Asti	0.20	0.60	0.33	0.17	0.60	1	1	0.92
38 San Telmo Merlot	0.60	0.80	1	0.83	1	1	1	1
39 Williams Selyem Chardonnay	1	1	1	1	0.60	0.80	1	0.83
40 The Foundry Syrah	1	0.80	1	0.92	1	0.80	1	0.92
	0.76	0.68	0.73	0.70	0.93	0.88	0.92	0.91

	Sweet				Wines			
	C	R	W	Total	C	R	W	Total
1 Louis Jadot Pommard	0.50	0.71	1	0.80	1	1	1	1
2 Nobilo "Icon" Sauvignon Blanc	0.50	0.43	0.33	0.40	1	1	1	1
3 Bonterra Chardonnay	0.50	0.71	0.56	0.60	1	1	1	1
4 Blue Mountain Icewine	0.50	1	1	0.90	1	1	1	1
5 Chateau Lafite Rothschild	0.50	1	1	0.90	1	1	1	1
6 Darting Riesling Kabinett	0.50	1	0.11	0.50	1	1	1	1
7 Frog's Leap Rutherford	0.50	1	1	0.90	1	1	1	1
8 Rizzi Barbaresco	0.50	1	0.78	0.80	1	1	1	1
9 Blackstone Pinot Noir	0.50	0.71	1	0.80	1	1	1	1
10 Gnemiz Tocai Friulano	0.50	0.71	0.11	0.40	1	1	1	1
11 Beringer White Zinfandel	0	0.71	0.78	0.60	1	1	1	1
12 Mount Nittany Bergwein	0	1	0.56	0.60	1	1	1	1
13 Louis Jadot Pouilly Fuisse	0.50	0.71	0.78	0.70	1	1	1	1
14 Taylor Lake Country Red	0.50	0.43	0.56	0.30	1	1	1	1
15 Villa Sandi Prosecco	0.50	0.43	0.33	0.20	1	1	1	1
16 Straccali Chianti	0.50	1	1	0.90	1	1	1	1
17 Barefoot Chardonnay	0.50	0.43	0.56	0.50	1	1	1	1
18 Conde de Vimioso	0.50	0.71	0.56	0.60	1	1	1	1
19 Veuve Clicquot Ponsardin Brut	0.50	0.43	0.56	0.50	1	1	1	1
20 Gloria Ferrer Chard. (Estate)	0.50	0.71	0.56	0.60	1	1	1	1
21 Bremerton Tamblyn	0.50	0.71	0.78	0.70	1	1	1	1
22 B.P. Rothschild Pinot Noir	0.50	0.71	1	0.80	1	1	1	1
23 Chateau Frank Brut	0.50	0.43	0.56	0.50	1	1	1	1
24 Leapfrogmilch	0.50	0.71	0.56	0.60	1	1	1	1
25 Gloria Ferrer Chard. (Reserve)	0.50	0.71	0.78	0.70	1	1	1	1
26 McWilliam's Shiraz	0.50	0.71	0.78	0.70	1	1	1	1
27 Blue Mountain Riesling	0	0.71	0.56	0.50	1	1	1	1
28 Vendange Merlot	0.50	0.71	1	0.80	1	1	1	1
29 Twin Fin Cabernet Sauvignon	0.50	1	0.78	0.80	1	1	1	1
30 B.P. Rothschild Cab. Sauv.	0.50	1	1	0.90	1	1	1	1
31 3 Bridges Botrytis Semillon	1	1	1	1	1	1	1	1
32 Gaja Ca'Marcanda Promis	0.50	1	0.33	0.60	1	1	1	1
33 de Ladoucette Pouilly Fume	0.50	0.71	0.78	0.70	1	1	1	1
34 Snoqualmie Sauvignon Blanc	0.50	0.43	0.33	0.40	1	1	1	1
35 Chaddsford Merican	0.50	0.71	1	0.80	1	1	1	1
36 Turkey Flat Butchers Block	0.50	0.71	0.78	0.70	1	1	1	1
37 Elmo Pio Asti	0.50	1	1	0.90	1	1	1	1
38 San Telmo Merlot	0.50	1	1	0.90	1	1	1	1
39 Williams Selyem Chardonnay	0.50	0.71	0.56	0.60	1	1	1	1
40 The Foundry Syrah	0.50	1	0.78	0.80	1	1	1	1
	0.48	0.76	0.71	0.67	1.00	1.00	1.00	1.00

APPENDIX T:

TYPICALITY RATINGS

The following tables present the mean (and standard deviation) typicality ratings for each wine-category pair, calculated by expert group and overall. To improve legibility, blank cells indicate averages or standard deviations of zero. I have reported the category-wide averages because the group differences may be informative. However, by themselves, the averages are not particularly meaningful because they depend on the specific set of wines being evaluated. Knowing that the mean typicality rating for the champagne category was lower than the mean for single varietal wines just tells us that there were fewer champagnes in the set. On the other hand, the fact that winemakers had a higher average than connoisseurs for the high end wines may speak to differences in their perception of the wines.

		Bordeaux Blend							
	Name	Connoisseur		Retailer		Winemaker		Total	
1	Louis Jadot Pommard	0.40	*(0.80)*	0.60	*(1.28)*	0.78	*(1.47)*	0.63	*(1.28)*
2	Nobilo "Icon" Sauvignon Blanc			0.90	*(1.81)*	0.11	*(0.31)*	0.42	*(1.26)*
3	Bonterra Chardonnay			0.50	*(1.50)*	0.11	*(0.31)*	0.25	*(1.01)*
4	Blue Mountain Icewine			0.60	*(1.80)*	0.11	*(0.31)*	0.29	*(1.21)*
5	Chateau Lafite Rothschild	7.00		6.80	*(0.60)*	6.67	*(0.94)*	6.79	*(0.71)*
6	Darting Riesling Kabinett					0.11	*(0.31)*	0.04	*(0.20)*
7	Frog's Leap Rutherford	5.00	*(2.61)*	5.90	*(1.30)*	5.11	*(2.28)*	5.42	*(2.06)*
8	Rizzi Barbaresco			0.50	*(0.92)*	0.56	*(1.26)*	0.42	*(1.00)*
9	Blackstone Pinot Noir			0.70	*(1.55)*	0.33	*(0.67)*	0.42	*(1.11)*
10	Gnemiz Tocai Friulano					0.11	*(0.31)*	0.04	*(0.20)*
11	Beringer White Zinfandel			0.10	*(0.30)*			0.04	*(0.20)*
12	Mount Nittany Bergwein					0.22	*(0.63)*	0.08	*(0.40)*
13	Louis Jadot Pouilly Fuisse			0.40	*(1.20)*	0.33	*(0.67)*	0.29	*(0.89)*
14	Taylor Lake Country Red			0.30	*(0.64)*	0.33	*(0.67)*	0.25	*(0.60)*
15	Villa Sandi Prosecco			0.10	*(0.30)*			0.04	*(0.20)*
16	Straccali Chianti			0.60	*(1.02)*	0.78	*(1.47)*	0.54	*(1.15)*
17	Barefoot Chardonnay			0.40	*(1.20)*	0.11	*(0.31)*	0.21	*(0.82)*
18	Conde de Vimioso	0.20	*(0.40)*	0.20	*(0.40)*	0.63	*(1.32)*	0.35	*(0.87)*
19	Veuve Clicquot Ponsardin Brut			0.40	*(1.20)*	0.11	*(0.31)*	0.21	*(0.82)*
20	Gloria Ferrer Chard. (Estate)			0.40	*(1.20)*	0.11	*(0.31)*	0.21	*(0.82)*
21	Bremerton Tamblyn	1.60	*(1.62)*	3.80	*(2.52)*	2.22	*(2.10)*	2.75	*(2.38)*
22	B.P. Rothschild Pinot Noir	1.20	*(2.40)*	0.70	*(1.55)*	0.89	*(1.66)*	0.88	*(1.81)*
23	Chateau Frank Brut					0.11	*(0.31)*	0.04	*(0.20)*
24	Leapfrogmilch			0.10	*(0.30)*	0.11	*(0.31)*	0.08	*(0.28)*
25	Gloria Ferrer Chard. (Reserve)			0.50	*(1.50)*	0.11	*(0.31)*	0.25	*(1.01)*
26	McWilliam's Shiraz			1.00	*(1.61)*	1.33	*(1.83)*	0.92	*(1.61)*
27	Blue Mountain Riesling			0.10	*(0.30)*	0.11	*(0.31)*	0.08	*(0.28)*
28	Vendange Merlot	0.60	*(1.20)*	1.80	*(2.56)*	1.22	*(1.69)*	1.33	*(2.07)*
29	Twin Fin Cabernet Sauvignon	0.60	*(1.20)*	2.40	*(2.42)*	1.22	*(1.69)*	1.58	*(2.08)*
30	B.P. Rothschild Cab. Sauv.	4.00	*(2.68)*	3.50	*(2.73)*	2.38	*(2.74)*	3.22	*(2.80)*
31	3 Bridges Botrytis Semillon			1.20	*(2.40)*	0.56	*(1.57)*	0.71	*(1.88)*
32	Gaja Ca'Marcanda Promis	0.20	*(0.40)*	1.56	*(2.41)*	0.67	*(1.56)*	0.91	*(1.89)*
33	de Ladoucette Pouilly Fume			1.20	*(1.89)*	0.44	*(0.96)*	0.67	*(1.43)*
34	Snoqualmie Sauvignon Blanc	1.20	*(2.40)*	1.30	*(2.61)*	0.11	*(0.31)*	0.83	*(2.09)*
35	Chaddsford Merican	2.20	*(2.86)*	2.40	*(2.58)*	5.56	*(1.83)*	3.54	*(2.86)*
36	Turkey Flat Butchers Block			1.00	*(1.41)*	1.33	*(1.83)*	0.92	*(1.53)*
37	Elmo Pio Asti								
38	San Telmo Merlot	0.60	*(1.20)*	2.20	*(2.44)*	1.22	*(1.69)*	1.50	*(2.06)*
39	Williams Selyem Chardonnay			0.40	*(1.20)*	0.11	*(0.31)*	0.21	*(0.82)*
40	The Foundry Syrah			0.40	*(0.92)*	0.44	*(1.26)*	0.33	*(0.99)*
		0.62	*(1.79)*	1.12	*(2.16)*	0.92	*(1.92)*	0.94	*(2.01)*

	Name	Champagne							
		Connoisseur		Retailer		Winemaker		Total	
1	Louis Jadot Pommard								
2	Nobilo "Icon" Sauvignon Blanc			0.38	*(0.99)*			0.14	*(0.64)*
3	Bonterra Chardonnay			0.38	*(0.99)*			0.14	*(0.64)*
4	Blue Mountain Icewine								
5	Chateau Lafite Rothschild								
6	Darting Riesling Kabinett								
7	Frog's Leap Rutherford								
8	Rizzi Barbaresco								
9	Blackstone Pinot Noir								
10	Gnemiz Tocai Friulano								
11	Beringer White Zinfandel								
12	Mount Nittany Bergwein								
13	Louis Jadot Pouilly Fuisse			0.38	*(0.99)*			0.14	*(0.64)*
14	Taylor Lake Country Red								
15	Villa Sandi Prosecco	2.00	*(2.12)*	1.75	*(2.38)*	1.56	*(1.64)*	1.71	*(2.05)*
16	Straccali Chianti								
17	Barefoot Chardonnay			0.38	*(0.99)*			0.14	*(0.64)*
18	Conde de Vimioso								
19	Veuve Clicquot Ponsardin Brut	7.00		7.00		7.00		7.00	
20	Gloria Ferrer Chard. (Estate)			0.38	*(0.99)*			0.14	*(0.64)*
21	Bremerton Tamblyn								
22	B.P. Rothschild Pinot Noir								
23	Chateau Frank Brut	2.50	*(2.06)*	2.63	*(2.96)*	4.44	*(2.75)*	3.38	*(2.87)*
24	Leapfrogmilch								
25	Gloria Ferrer Chard. (Reserve)			0.38	*(0.99)*			0.14	*(0.64)*
26	McWilliam's Shiraz								
27	Blue Mountain Riesling								
28	Vendange Merlot								
29	Twin Fin Cabernet Sauvignon								
30	B.P. Rothschild Cab. Sauv.								
31	3 Bridges Botrytis Semillon								
32	Gaja Ca'Marcanda Promis					0.22	*(0.63)*	0.10	*(0.43)*
33	de Ladoucette Pouilly Fume			0.38	*(0.99)*			0.14	*(0.64)*
34	Snoqualmie Sauvignon Blanc			0.38	*(0.99)*			0.14	*(0.64)*
35	Chaddsford Merican								
36	Turkey Flat Butchers Block								
37	Elmo Pio Asti	2.75	*(1.79)*	1.50	*(2.35)*	1.33	*(1.70)*	1.67	*(2.05)*
38	San Telmo Merlot								
39	Williams Selyem Chardonnay			0.38	*(0.99)*			0.14	*(0.64)*
40	The Foundry Syrah								
		0.36	*(1.36)*	0.41	*(1.45)*	0.36	*(1.43)*	0.38	*(1.42)*

	Name	High End						Total	
		Connoisseur		Retailer		Winemaker		Total	
1	Louis Jadot Pommard	4.50	*(1.80)*	5.25	*(1.85)*	6.00	*(1.25)*	5.43	*(1.71)*
2	Nobilo "Icon" Sauvignon Blanc	2.00	*(2.12)*	2.38	*(1.65)*	3.22	*(2.10)*	2.67	*(2.01)*
3	Bonterra Chardonnay	1.00	*(1.73)*	1.75	*(1.20)*	3.00	*(1.94)*	2.14	*(1.83)*
4	Blue Mountain Icewine			3.63	*(2.29)*	4.38	*(2.39)*	3.20	*(2.66)*
5	Chateau Lafite Rothschild	7.00		7.00		6.89	*(0.31)*	6.95	*(0.21)*
6	Darting Riesling Kabinett	2.00	*(1.22)*	1.88	*(1.62)*	3.78	*(2.48)*	2.71	*(2.19)*
7	Frog's Leap Rutherford	2.75	*(1.64)*	6.00	*(1.22)*	5.00	*(2.05)*	4.95	*(2.06)*
8	Rizzi Barbaresco	2.00	*(2.12)*	5.38	*(1.32)*	4.44	*(2.01)*	4.33	*(2.17)*
9	Blackstone Pinot Noir	1.00	*(1.73)*	1.63	*(1.11)*	3.56	*(2.67)*	2.33	*(2.30)*
10	Gnemiz Tocai Friulano			3.00	*(2.00)*	2.00	*(1.91)*	2.06	*(2.08)*
11	Beringer White Zinfandel			0.25	*(0.43)*	0.78	*(1.31)*	0.43	*(0.95)*
12	Mount Nittany Bergwein			0.43	*(0.49)*	1.44	*(1.71)*	0.80	*(1.33)*
13	Louis Jadot Pouilly Fuisse	4.00	*(1.41)*	4.63	*(1.80)*	4.22	*(2.70)*	4.33	*(2.19)*
14	Taylor Lake Country Red			0.88	*(1.96)*	0.67	*(1.25)*	0.62	*(1.50)*
15	Villa Sandi Prosecco	0.50	*(0.87)*	0.86	*(1.12)*	2.25	*(1.56)*	1.37	*(1.49)*
16	Straccali Chianti	2.00	*(1.41)*	2.00	*(2.18)*	3.89	*(2.33)*	2.85	*(2.35)*
17	Barefoot Chardonnay			1.13	*(1.36)*	2.22	*(2.20)*	1.38	*(1.86)*
18	Conde de Vimioso			1.00	*(1.41)*	2.60	*(1.62)*	1.33	*(1.66)*
19	Veuve Clicquot Ponsardin Brut	5.75	*(1.30)*	6.50	*(0.71)*	6.67	*(0.67)*	6.43	*(0.90)*
20	Gloria Ferrer Chard. (Estate)	2.25	*(1.30)*	3.50	*(1.80)*	4.00	*(2.58)*	3.48	*(2.20)*
21	Bremerton Tamblyn	1.75	*(2.05)*	2.71	*(2.12)*	3.67	*(2.36)*	2.95	*(2.33)*
22	B.P. Rothschild Pinot Noir	3.25	*(2.17)*	0.75	*(1.30)*	3.22	*(2.44)*	2.29	*(2.35)*
23	Chateau Frank Brut			1.38	*(1.22)*	4.11	*(2.64)*	2.29	*(2.51)*
24	Leapfrogmilch	1.00	*(1.73)*	2.25	*(1.39)*	1.22	*(1.40)*	1.57	*(1.56)*
25	Gloria Ferrer Chard. (Reserve)	2.50	*(1.50)*	3.38	*(2.29)*	4.67	*(2.05)*	3.76	*(2.22)*
26	McWilliam's Shiraz	2.25	*(1.92)*	4.63	*(1.22)*	5.25	*(1.39)*	4.40	*(1.83)*
27	Blue Mountain Riesling			0.88	*(1.05)*	2.44	*(1.71)*	1.38	*(1.62)*
28	Vendange Merlot	0.50	*(0.87)*	0.25	*(0.43)*	2.33	*(2.26)*	1.19	*(1.84)*
29	Twin Fin Cabernet Sauvignon			0.75	*(1.09)*	1.33	*(1.33)*	0.86	*(1.21)*
30	B.P. Rothschild Cab. Sauv.	3.25	*(2.49)*	0.63	*(0.99)*	3.22	*(2.44)*	2.24	*(2.39)*
31	3 Bridges Botrytis Semillon	1.00	*(1.22)*	3.13	*(2.26)*	5.25	*(2.22)*	3.55	*(2.62)*
32	Gaja Ca'Marcanda Promis	2.50	*(2.87)*	6.25	*(1.39)*	5.40	*(1.62)*	5.12	*(2.42)*
33	de Ladoucette Pouilly Fume	5.00	*(1.41)*	4.50	*(1.50)*	5.56	*(1.77)*	5.05	*(1.69)*
34	Snoqualmie Sauvignon Blanc	1.50	*(1.50)*	2.13	*(1.36)*	2.89	*(1.73)*	2.33	*(1.64)*
35	Chaddsford Merican			1.50	*(2.24)*	4.22	*(1.87)*	2.38	*(2.50)*
36	Turkey Flat Butchers Block	2.00	*(1.87)*	3.00	*(2.39)*	3.00	*(2.16)*	2.80	*(2.23)*
37	Elmo Pio Asti			0.88	*(1.05)*	1.78	*(1.81)*	1.10	*(1.51)*
38	San Telmo Merlot	1.33	*(1.89)*	0.86	*(1.12)*	2.00	*(1.41)*	1.44	*(1.50)*
39	Williams Selyem Chardonnay	5.50	*(1.12)*	4.88	*(2.09)*	5.78	*(1.23)*	5.38	*(1.65)*
40	The Foundry Syrah	2.00	*(2.16)*	3.57	*(1.92)*	3.43	*(1.68)*	3.24	*(1.96)*
		1.81	*(2.31)*	2.70	*(2.48)*	3.54	*(2.51)*	2.89	*(2.54)*

	Name	Italian						Total	
		Connoisseur		Retailer		Winemaker			
1	Louis Jadot Pommard			0.40	*(0.92)*			0.17	*(0.62)*
2	Nobilo "Icon" Sauvignon Blanc			0.60	*(1.50)*	0.33	*(0.94)*	0.38	*(1.15)*
3	Bonterra Chardonnay			0.60	*(1.20)*			0.25	*(0.83)*
4	Blue Mountain Icewine	0.80	*(1.60)*	0.60	*(1.50)*			0.42	*(1.26)*
5	Chateau Lafite Rothschild			0.30	*(0.90)*			0.13	*(0.60)*
6	Darting Riesling Kabinett			0.50	*(1.50)*			0.21	*(1.00)*
7	Frog's Leap Rutherford			0.30	*(0.90)*			0.13	*(0.60)*
8	Rizzi Barbaresco	5.60	*(2.80)*	6.90	*(0.30)*	6.89	*(0.31)*	6.63	*(1.41)*
9	Blackstone Pinot Noir			0.50	*(1.02)*			0.21	*(0.71)*
10	Gnemiz Tocai Friulano	6.20	*(1.17)*	6.90	*(0.30)*	6.00	*(2.21)*	6.42	*(1.53)*
11	Beringer White Zinfandel			0.10	*(0.30)*	0.33	*(0.94)*	0.17	*(0.62)*
12	Mount Nittany Bergwein			0.60	*(1.50)*	0.33	*(0.94)*	0.38	*(1.15)*
13	Louis Jadot Pouilly Fuisse			0.30	*(0.90)*			0.13	*(0.60)*
14	Taylor Lake Country Red			0.10	*(0.30)*	0.22	*(0.63)*	0.13	*(0.44)*
15	Villa Sandi Prosecco	5.40	*(1.36)*	6.30	*(1.42)*	5.89	*(2.18)*	5.96	*(1.77)*
16	Straccali Chianti	6.80	*(0.40)*	6.60	*(0.92)*	7.00		6.79	*(0.64)*
17	Barefoot Chardonnay			0.50	*(1.02)*			0.21	*(0.71)*
18	Conde de Vimioso			2.00	*(2.90)*	1.44	*(2.50)*	1.38	*(2.53)*
19	Veuve Clicquot Ponsardin Brut			0.20	*(0.40)*			0.08	*(0.28)*
20	Gloria Ferrer Chard. (Estate)			0.60	*(1.20)*			0.25	*(0.83)*
21	Bremerton Tamblyn			0.20	*(0.60)*			0.08	*(0.40)*
22	B.P. Rothschild Pinot Noir			0.50	*(1.02)*			0.21	*(0.71)*
23	Chateau Frank Brut			0.40	*(0.92)*			0.17	*(0.62)*
24	Leapfrogmilch			0.50	*(1.50)*			0.21	*(1.00)*
25	Gloria Ferrer Chard. (Reserve)			0.60	*(1.20)*			0.25	*(0.83)*
26	McWilliam's Shiraz			0.10	*(0.30)*			0.04	*(0.20)*
27	Blue Mountain Riesling	0.80	*(1.60)*	0.20	*(0.60)*			0.25	*(0.88)*
28	Vendange Merlot			0.20	*(0.60)*			0.08	*(0.40)*
29	Twin Fin Cabernet Sauvignon			0.50	*(1.02)*			0.21	*(0.71)*
30	B.P. Rothschild Cab. Sauv.			0.30	*(0.90)*			0.13	*(0.60)*
31	3 Bridges Botrytis Semillon			1.00	*(2.00)*			0.42	*(1.38)*
32	Gaja Ca'Marcanda Promis	6.80	*(0.40)*	5.80	*(2.09)*	6.67	*(0.67)*	6.33	*(1.49)*
33	de Ladoucette Pouilly Fume			0.30	*(0.64)*			0.13	*(0.44)*
34	Snoqualmie Sauvignon Blanc			0.60	*(1.50)*			0.25	*(1.01)*
35	Chaddsford Merican			0.10	*(0.30)*	0.22	*(0.63)*	0.13	*(0.44)*
36	Turkey Flat Butchers Block			0.20	*(0.60)*			0.08	*(0.40)*
37	Elmo Pio Asti	5.60	*(1.50)*	6.30	*(1.42)*	6.89	*(0.31)*	6.38	*(1.25)*
38	San Telmo Merlot			0.90	*(1.81)*			0.38	*(1.25)*
39	Williams Selyem Chardonnay			0.60	*(1.20)*			0.25	*(0.83)*
40	The Foundry Syrah			0.10	*(0.30)*			0.04	*(0.20)*
		0.95	*(2.27)*	1.36	*(2.48)*	1.06	*(2.43)*	1.16	*(2.43)*

	Name	Light White						Total	
		Connoisseur		Retailer		Winemaker		Total	
1	Louis Jadot Pommard			0.10	*(0.30)*			0.04	*(0.20)*
2	Nobilo "Icon" Sauvignon Blanc	4.20	*(2.48)*	5.70	*(1.27)*	5.11	*(2.77)*	5.17	*(2.27)*
3	Bonterra Chardonnay	3.20	*(1.94)*	2.90	*(2.17)*	3.56	*(2.45)*	3.21	*(2.25)*
4	Blue Mountain Icewine	2.00	*(1.41)*	2.00	*(2.24)*	1.67	*(1.89)*	1.88	*(1.96)*
5	Chateau Lafite Rothschild			0.40	*(1.20)*			0.17	*(0.80)*
6	Darting Riesling Kabinett	5.80	*(1.17)*	5.00	*(2.61)*	6.56	*(0.68)*	5.75	*(1.94)*
7	Frog's Leap Rutherford			0.10	*(0.30)*			0.04	*(0.20)*
8	Rizzi Barbaresco								
9	Blackstone Pinot Noir			0.10	*(0.30)*	0.78	*(2.20)*	0.33	*(1.40)*
10	Gnemiz Tocai Friulano	2.25	*(2.49)*	4.60	*(2.15)*	4.89	*(2.88)*	4.30	*(2.69)*
11	Beringer White Zinfandel	3.40	*(2.73)*	1.00	*(1.90)*	2.78	*(2.86)*	2.17	*(2.67)*
12	Mount Nittany Bergwein	3.50	*(2.06)*	5.11	*(2.13)*	4.44	*(2.79)*	4.55	*(2.48)*
13	Louis Jadot Pouilly Fuisse	2.80	*(2.79)*	3.10	*(2.21)*	3.44	*(2.75)*	3.17	*(2.56)*
14	Taylor Lake Country Red					0.78	*(2.20)*	0.29	*(1.40)*
15	Villa Sandi Prosecco	2.80	*(1.72)*	5.20	*(1.66)*	4.44	*(2.59)*	4.42	*(2.25)*
16	Straccali Chianti								
17	Barefoot Chardonnay	4.80	*(1.47)*	3.10	*(1.97)*	3.44	*(2.59)*	3.58	*(2.23)*
18	Conde de Vimioso								
19	Veuve Clicquot Ponsardin Brut	2.60	*(2.42)*	3.30	*(2.76)*	3.67	*(2.71)*	3.29	*(2.70)*
20	Gloria Ferrer Chard. (Estate)	3.40	*(2.06)*	2.70	*(2.19)*	2.89	*(2.38)*	2.92	*(2.25)*
21	Bremerton Tamblyn								
22	B.P. Rothschild Pinot Noir			0.10	*(0.30)*			0.04	*(0.20)*
23	Chateau Frank Brut	2.00	*(1.41)*	3.70	*(2.65)*	4.44	*(2.54)*	3.63	*(2.56)*
24	Leapfrogmilch	4.80	*(1.33)*	4.70	*(2.00)*	4.56	*(2.59)*	4.67	*(2.13)*
25	Gloria Ferrer Chard. (Reserve)	2.80	*(1.72)*	2.50	*(2.06)*	2.44	*(1.95)*	2.54	*(1.96)*
26	McWilliam's Shiraz								
27	Blue Mountain Riesling	5.40	*(1.02)*	5.30	*(2.10)*	5.11	*(2.38)*	5.25	*(2.05)*
28	Vendange Merlot			0.10	*(0.30)*			0.04	*(0.20)*
29	Twin Fin Cabernet Sauvignon			0.10	*(0.30)*			0.04	*(0.20)*
30	B.P. Rothschild Cab. Sauv.								
31	3 Bridges Botrytis Semillon	2.00	*(1.41)*	2.70	*(2.87)*	2.56	*(2.41)*	2.50	*(2.47)*
32	Gaja Ca'Marcanda Promis								
33	de Ladoucette Pouilly Fume	4.20	*(2.79)*	4.20	*(2.36)*	4.56	*(2.54)*	4.33	*(2.53)*
34	Snoqualmie Sauvignon Blanc	5.40	*(1.85)*	5.80	*(0.87)*	4.78	*(2.70)*	5.33	*(1.99)*
35	Chaddsford Merican								
36	Turkey Flat Butchers Block			0.10	*(0.30)*			0.04	*(0.20)*
37	Elmo Pio Asti	3.40	*(2.33)*	4.70	*(2.33)*	4.56	*(2.63)*	4.38	*(2.50)*
38	San Telmo Merlot								
39	Williams Selyem Chardonnay	3.20	*(2.32)*	2.10	*(2.12)*	2.67	*(2.36)*	2.54	*(2.29)*
40	The Foundry Syrah			0.10	*(0.30)*			0.04	*(0.20)*
		1.84	*(2.43)*	2.01	*(2.63)*	2.10	*(2.81)*	2.01	*(2.66)*

	Name	Old World							
		Connoisseur		Retailer		Winemaker		Total	
1	Louis Jadot Pommard	5.60	*(2.80)*	6.50	*(1.02)*	7.00		6.50	*(1.53)*
2	Nobilo "Icon" Sauvignon Blanc			0.60	*(1.28)*	0.44	*(1.26)*	0.42	*(1.15)*
3	Bonterra Chardonnay			0.60	*(1.28)*	0.44	*(0.83)*	0.42	*(1.00)*
4	Blue Mountain Icewine			0.50	*(1.20)*	0.67	*(1.56)*	0.46	*(1.26)*
5	Chateau Lafite Rothschild	7.00		6.80	*(0.60)*	7.00		6.92	*(0.40)*
6	Darting Riesling Kabinett	3.80	*(3.12)*	5.80	*(2.09)*	6.11	*(2.18)*	5.50	*(2.53)*
7	Frog's Leap Rutherford	1.40	*(2.80)*	1.00	*(2.05)*	1.33	*(2.54)*	1.21	*(2.41)*
8	Rizzi Barbaresco	5.20	*(2.64)*	6.80	*(0.40)*	6.89	*(0.31)*	6.50	*(1.41)*
9	Blackstone Pinot Noir			0.40	*(0.92)*	0.44	*(1.26)*	0.33	*(0.99)*
10	Gnemiz Tocai Friulano	4.60	*(2.58)*	5.90	*(2.02)*	6.67	*(0.67)*	5.92	*(1.96)*
11	Beringer White Zinfandel			0.20	*(0.40)*			0.08	*(0.28)*
12	Mount Nittany Bergwein			0.20	*(0.40)*			0.08	*(0.28)*
13	Louis Jadot Pouilly Fuisse	5.60	*(2.80)*	6.60	*(0.80)*	7.00		6.54	*(1.47)*
14	Taylor Lake Country Red			0.20	*(0.40)*			0.08	*(0.28)*
15	Villa Sandi Prosecco	4.00	*(2.45)*	6.60	*(0.92)*	6.33	*(1.33)*	5.96	*(1.81)*
16	Straccali Chianti	5.60	*(2.80)*	6.50	*(0.92)*	6.89	*(0.31)*	6.46	*(1.50)*
17	Barefoot Chardonnay			0.60	*(1.28)*	0.22	*(0.63)*	0.33	*(0.94)*
18	Conde de Vimioso	4.60	*(2.50)*	6.20	*(1.54)*	5.78	*(2.20)*	5.71	*(2.11)*
19	Veuve Clicquot Ponsardin Brut	5.00	*(2.61)*	6.30	*(1.79)*	7.00		6.29	*(1.81)*
20	Gloria Ferrer Chard. (Estate)			0.60	*(1.28)*	0.56	*(1.26)*	0.46	*(1.15)*
21	Bremerton Tamblyn			0.70	*(1.55)*	0.56	*(1.57)*	0.50	*(1.41)*
22	B.P. Rothschild Pinot Noir	4.60	*(2.50)*	5.00	*(2.45)*	5.33	*(2.16)*	5.04	*(2.37)*
23	Chateau Frank Brut			0.50	*(1.20)*	0.78	*(1.62)*	0.50	*(1.29)*
24	Leapfrogmilch	1.40	*(2.80)*	0.50	*(1.20)*	0.11	*(0.31)*	0.54	*(1.58)*
25	Gloria Ferrer Chard. (Reserve)			0.60	*(1.28)*	0.44	*(0.83)*	0.42	*(1.00)*
26	McWilliam's Shiraz			1.20	*(2.44)*	0.44	*(1.26)*	0.70	*(1.85)*
27	Blue Mountain Riesling			0.70	*(1.79)*	0.22	*(0.63)*	0.38	*(1.25)*
28	Vendange Merlot			0.60	*(1.28)*	0.44	*(1.26)*	0.42	*(1.15)*
29	Twin Fin Cabernet Sauvignon			0.30	*(0.64)*	0.44	*(1.26)*	0.29	*(0.89)*
30	B.P. Rothschild Cab. Sauv.	4.80	*(2.64)*	5.30	*(2.24)*	5.44	*(2.17)*	5.25	*(2.31)*
31	3 Bridges Botrytis Semillon			0.60	*(1.28)*	0.67	*(1.56)*	0.50	*(1.29)*
32	Gaja Ca'Marcanda Promis	5.20	*(2.64)*	6.22	*(1.23)*	5.89	*(2.18)*	5.87	*(2.03)*
33	de Ladoucette Pouilly Fume	5.60	*(2.80)*	6.60	*(0.80)*	7.00		6.54	*(1.47)*
34	Snoqualmie Sauvignon Blanc			0.70	*(1.55)*	0.67	*(1.33)*	0.54	*(1.32)*
35	Chaddsford Merican			0.20	*(0.40)*	1.00	*(1.89)*	0.46	*(1.26)*
36	Turkey Flat Butchers Block			0.60	*(1.28)*	0.56	*(1.57)*	0.46	*(1.29)*
37	Elmo Pio Asti	4.00	*(2.45)*	6.10	*(1.45)*	6.00	*(1.89)*	5.63	*(2.04)*
38	San Telmo Merlot	0.60	*(1.20)*	0.60	*(1.50)*	0.56	*(1.26)*	0.58	*(1.35)*
39	Williams Selyem Chardonnay			0.60	*(1.28)*	0.44	*(1.26)*	0.42	*(1.15)*
40	The Foundry Syrah	0.60	*(1.20)*	0.60	*(1.28)*	0.56	*(1.57)*	0.58	*(1.38)*
		1.99	*(2.97)*	2.67	*(3.09)*	2.71	*(3.20)*	2.54	*(3.12)*

	Name	Party						Total	
		Connoisseur		Retailer		Winemaker			
1	Louis Jadot Pommard	2.40	*(1.85)*	1.30	*(1.68)*	1.44	*(2.17)*	1.58	*(1.96)*
2	Nobilo "Icon" Sauvignon Blanc	3.60	*(2.33)*	3.80	*(2.18)*	2.44	*(2.17)*	3.25	*(2.30)*
3	Bonterra Chardonnay	4.00	*(2.10)*	3.30	*(2.49)*	2.78	*(2.78)*	3.25	*(2.57)*
4	Blue Mountain Icewine	1.00	*(0.89)*	1.10	*(1.45)*	2.67	*(2.67)*	1.67	*(2.07)*
5	Chateau Lafite Rothschild	1.20	*(0.98)*	0.90	*(1.45)*	1.44	*(2.17)*	1.17	*(1.70)*
6	Darting Riesling Kabinett	3.20	*(2.32)*	1.20	*(1.25)*	4.22	*(2.20)*	2.75	*(2.33)*
7	Frog's Leap Rutherford	3.00	*(1.67)*	1.50	*(1.63)*	1.44	*(2.11)*	1.79	*(1.94)*
8	Rizzi Barbaresco	2.00	*(1.22)*	1.30	*(1.68)*	2.00	*(2.31)*	1.70	*(1.92)*
9	Blackstone Pinot Noir	3.40	*(1.85)*	4.90	*(1.37)*	1.78	*(2.20)*	3.42	*(2.29)*
10	Gnemiz Tocai Friulano	4.00	*(2.74)*	3.22	*(2.53)*	3.50	*(2.40)*	3.48	*(2.54)*
11	Beringer White Zinfandel	5.40	*(2.33)*	5.20	*(2.23)*	6.22	*(1.03)*	5.63	*(1.95)*
12	Mount Nittany Bergwein	3.00	*(1.79)*	2.50	*(2.84)*	5.44	*(1.71)*	3.71	*(2.64)*
13	Louis Jadot Pouilly Fuisse	3.00	*(2.00)*	2.50	*(2.11)*	2.11	*(2.13)*	2.46	*(2.12)*
14	Taylor Lake Country Red	3.00	*(2.10)*	3.10	*(2.98)*	5.89	*(1.45)*	4.13	*(2.70)*
15	Villa Sandi Prosecco	5.60	*(0.80)*	5.30	*(2.24)*	4.89	*(2.33)*	5.21	*(2.08)*
16	Straccali Chianti	3.40	*(2.06)*	4.70	*(1.55)*	1.89	*(2.18)*	3.38	*(2.29)*
17	Barefoot Chardonnay	6.00	*(1.26)*	5.70	*(1.49)*	3.67	*(2.62)*	5.00	*(2.22)*
18	Conde de Vimioso	2.60	*(2.06)*	2.80	*(2.52)*	2.11	*(2.60)*	2.50	*(2.48)*
19	Veuve Clicquot Ponsardin Brut	5.00	*(1.79)*	4.80	*(2.44)*	3.33	*(2.36)*	4.29	*(2.41)*
20	Gloria Ferrer Chard. (Estate)	3.80	*(2.04)*	2.30	*(2.05)*	2.33	*(2.40)*	2.63	*(2.27)*
21	Bremerton Tamblyn	2.80	*(2.14)*	2.22	*(2.20)*	1.67	*(2.21)*	2.13	*(2.23)*
22	B.P. Rothschild Pinot Noir	2.40	*(1.62)*	4.20	*(1.94)*	2.33	*(2.16)*	3.13	*(2.17)*
23	Chateau Frank Brut	5.80	*(0.75)*	4.20	*(2.99)*	3.33	*(2.36)*	4.21	*(2.60)*
24	Leapfrogmilch	4.80	*(1.83)*	2.70	*(2.28)*	5.67	*(1.63)*	4.25	*(2.38)*
25	Gloria Ferrer Chard. (Reserve)	2.80	*(1.72)*	2.50	*(2.46)*	2.11	*(2.47)*	2.42	*(2.34)*
26	McWilliam's Shiraz	3.00	*(1.67)*	2.20	*(2.18)*	1.56	*(2.17)*	2.13	*(2.15)*
27	Blue Mountain Riesling	2.60	*(2.06)*	2.60	*(2.54)*	5.22	*(1.47)*	3.58	*(2.45)*
28	Vendange Merlot	4.20	*(2.23)*	5.10	*(2.12)*	2.56	*(2.50)*	3.96	*(2.56)*
29	Twin Fin Cabernet Sauvignon	4.20	*(2.40)*	5.70	*(1.19)*	3.56	*(2.54)*	4.58	*(2.27)*
30	B.P. Rothschild Cab. Sauv.	3.00	*(2.45)*	4.80	*(1.72)*	2.00	*(2.58)*	3.38	*(2.56)*
31	3 Bridges Botrytis Semillon	0.60	*(0.49)*	1.10	*(1.45)*	2.56	*(2.71)*	1.54	*(2.08)*
32	Gaja Ca'Marcanda Promis	2.40	*(2.06)*	1.11	*(1.66)*	1.67	*(2.26)*	1.61	*(2.06)*
33	de Ladoucette Pouilly Fume	2.80	*(2.14)*	1.40	*(1.50)*	2.44	*(2.17)*	2.08	*(2.00)*
34	Snoqualmie Sauvignon Blanc	3.40	*(2.15)*	2.70	*(2.41)*	2.67	*(2.26)*	2.83	*(2.32)*
35	Chaddsford Merican	2.20	*(1.60)*	0.44	*(0.68)*	1.56	*(2.27)*	1.26	*(1.80)*
36	Turkey Flat Butchers Block	3.60	*(2.06)*	3.50	*(2.58)*	1.89	*(2.38)*	2.92	*(2.53)*
37	Elmo Pio Asti	6.40	*(0.80)*	5.10	*(2.26)*	5.11	*(2.42)*	5.38	*(2.18)*
38	San Telmo Merlot	4.00	*(2.19)*	3.50	*(2.73)*	2.33	*(2.36)*	3.17	*(2.58)*
39	Williams Selyem Chardonnay	2.20	*(1.72)*	1.50	*(1.80)*	2.11	*(2.38)*	1.88	*(2.05)*
40	The Foundry Syrah	3.00	*(2.19)*	1.60	*(1.80)*	2.00	*(2.11)*	2.04	*(2.07)*
		3.37	*(2.29)*	3.00	*(2.58)*	2.90	*(2.64)*	3.04	*(2.55)*

	Name	Sauvignon Blanc							
		Connoisseur		Retailer		Winemaker		Total	
1	Louis Jadot Pommard								
2	Nobilo "Icon" Sauvignon Blanc	7.00		6.60	*(0.80)*	7.00		6.83	*(0.55)*
3	Bonterra Chardonnay			0.10	*(0.30)*	0.56	*(1.26)*	0.25	*(0.83)*
4	Blue Mountain Icewine			0.50	*(1.20)*	0.44	*(0.96)*	0.38	*(0.99)*
5	Chateau Lafite Rothschild								
6	Darting Riesling Kabinett			0.10	*(0.30)*	0.44	*(0.83)*	0.21	*(0.58)*
7	Frog's Leap Rutherford								
8	Rizzi Barbaresco								
9	Blackstone Pinot Noir								
10	Gnemiz Tocai Friulano			0.40	*(0.80)*	0.78	*(1.62)*	0.50	*(1.20)*
11	Beringer White Zinfandel					0.22	*(0.42)*	0.08	*(0.28)*
12	Mount Nittany Bergwein			0.60	*(1.28)*	0.33	*(0.67)*	0.38	*(0.95)*
13	Louis Jadot Pouilly Fuisse	2.60	*(3.20)*	0.60	*(1.28)*	2.78	*(2.86)*	1.83	*(2.64)*
14	Taylor Lake Country Red								
15	Villa Sandi Prosecco			0.70	*(1.79)*	0.44	*(0.96)*	0.48	*(1.35)*
16	Straccali Chianti								
17	Barefoot Chardonnay			0.10	*(0.30)*	0.56	*(1.26)*	0.25	*(0.83)*
18	Conde de Vimioso								
19	Veuve Clicquot Ponsardin Brut			0.60	*(1.80)*	0.33	*(0.67)*	0.38	*(1.25)*
20	Gloria Ferrer Chard. (Estate)			0.10	*(0.30)*	0.56	*(1.26)*	0.25	*(0.83)*
21	Bremerton Tamblyn								
22	B.P. Rothschild Pinot Noir								
23	Chateau Frank Brut			0.70	*(1.79)*	0.33	*(0.67)*	0.42	*(1.26)*
24	Leapfrogmilch			0.20	*(0.40)*	0.33	*(0.67)*	0.21	*(0.50)*
25	Gloria Ferrer Chard. (Reserve)			0.10	*(0.30)*	0.56	*(1.26)*	0.25	*(0.83)*
26	McWilliam's Shiraz								
27	Blue Mountain Riesling			0.50	*(1.20)*	0.44	*(0.83)*	0.38	*(0.95)*
28	Vendange Merlot								
29	Twin Fin Cabernet Sauvignon								
30	B.P. Rothschild Cab. Sauv.								
31	3 Bridges Botrytis Semillon			0.60	*(1.28)*	0.44	*(0.68)*	0.42	*(0.95)*
32	Gaja Ca'Marcanda Promis								
33	de Ladoucette Pouilly Fume	6.80	*(0.40)*	4.80	*(2.96)*	6.67	*(0.94)*	5.92	*(2.22)*
34	Snoqualmie Sauvignon Blanc	7.00		6.50	*(0.81)*	6.89	*(0.31)*	6.75	*(0.60)*
35	Chaddsford Merican								
36	Turkey Flat Butchers Block								
37	Elmo Pio Asti			0.50	*(1.20)*	0.33	*(0.67)*	0.33	*(0.90)*
38	San Telmo Merlot								
39	Williams Selyem Chardonnay			0.10	*(0.30)*	0.56	*(1.26)*	0.25	*(0.83)*
40	The Foundry Syrah								
		0.59	*(1.93)*	0.61	*(1.79)*	0.78	*(1.96)*	0.67	*(1.89)*

		Single Varietal							
	Name	Connoisseur		Retailer		Winemaker		Total	
1	Louis Jadot Pommard	5.00	*(2.92)*	5.30	*(2.41)*	5.78	*(2.10)*	5.43	*(2.41)*
2	Nobilo "Icon" Sauvignon Blanc	6.60	*(0.80)*	6.11	*(1.10)*	6.78	*(0.63)*	6.48	*(0.93)*
3	Bonterra Chardonnay	6.40	*(1.20)*	6.44	*(0.83)*	6.89	*(0.31)*	6.61	*(0.82)*
4	Blue Mountain Icewine	5.50	*(2.60)*	5.50	*(2.33)*	6.44	*(1.26)*	5.87	*(2.09)*
5	Chateau Lafite Rothschild	1.40	*(2.80)*	2.20	*(2.99)*	0.67	*(1.05)*	1.46	*(2.50)*
6	Darting Riesling Kabinett	7.00		5.60	*(1.91)*	6.67	*(0.94)*	6.26	*(1.51)*
7	Frog's Leap Rutherford			1.40	*(2.24)*	0.67	*(1.05)*	0.83	*(1.67)*
8	Rizzi Barbaresco	4.80	*(2.56)*	5.78	*(2.15)*	5.22	*(2.86)*	5.35	*(2.56)*
9	Blackstone Pinot Noir	6.40	*(1.20)*	5.44	*(2.17)*	6.89	*(0.31)*	6.22	*(1.61)*
10	Gnemiz Tocai Friulano	1.75	*(3.03)*	5.67	*(1.70)*	5.83	*(2.61)*	4.89	*(2.85)*
11	Beringer White Zinfandel	6.20	*(1.17)*	3.70	*(2.87)*	5.89	*(2.28)*	5.04	*(2.64)*
12	Mount Nittany Bergwein	1.40	*(2.80)*	1.30	*(2.15)*	1.00	*(2.21)*	1.21	*(2.33)*
13	Louis Jadot Pouilly Fuisse	5.50	*(1.50)*	6.00	*(1.55)*	5.89	*(1.66)*	5.87	*(1.60)*
14	Taylor Lake Country Red			1.70	*(2.72)*			0.74	*(1.98)*
15	Villa Sandi Prosecco	3.50	*(3.50)*	4.22	*(2.70)*	2.88	*(2.93)*	3.57	*(3.02)*
16	Straccali Chianti	4.20	*(3.43)*	3.60	*(3.07)*	0.78	*(1.13)*	2.67	*(3.01)*
17	Barefoot Chardonnay	6.40	*(1.20)*	6.00	*(1.25)*	7.00		6.48	*(1.06)*
18	Conde de Vimioso			1.13	*(1.36)*	1.00	*(1.60)*	0.89	*(1.41)*
19	Veuve Clicquot Ponsardin Brut	2.80	*(3.43)*	1.10	*(1.76)*	1.78	*(2.44)*	1.71	*(2.52)*
20	Gloria Ferrer Chard. (Estate)	6.40	*(1.20)*	6.44	*(0.83)*	7.00		6.65	*(0.81)*
21	Bremerton Tamblyn			1.00	*(1.67)*	0.56	*(0.83)*	0.63	*(1.25)*
22	B.P. Rothschild Pinot Noir	6.60	*(0.80)*	5.40	*(2.15)*	6.78	*(0.42)*	6.17	*(1.60)*
23	Chateau Frank Brut			0.60	*(1.02)*	1.22	*(2.25)*	0.71	*(1.59)*
24	Leapfrogmilch			0.60	*(1.02)*	0.22	*(0.63)*	0.33	*(0.80)*
25	Gloria Ferrer Chard. (Reserve)	6.40	*(1.20)*	6.67	*(0.67)*	7.00		6.74	*(0.74)*
26	McWilliam's Shiraz	6.20	*(0.98)*	5.89	*(2.13)*	6.89	*(0.31)*	6.35	*(1.49)*
27	Blue Mountain Riesling	6.40	*(1.20)*	6.11	*(1.10)*	6.89	*(0.31)*	6.48	*(0.97)*
28	Vendange Merlot	6.40	*(1.20)*	4.60	*(2.65)*	6.89	*(0.31)*	5.83	*(2.09)*
29	Twin Fin Cabernet Sauvignon	5.00	*(2.76)*	5.33	*(2.21)*	6.78	*(0.63)*	5.83	*(2.08)*
30	B.P. Rothschild Cab. Sauv.	5.60	*(2.80)*	5.40	*(2.15)*	6.78	*(0.42)*	5.96	*(2.01)*
31	3 Bridges Botrytis Semillon	5.60	*(2.33)*	6.50	*(0.67)*	6.44	*(1.26)*	6.29	*(1.43)*
32	Gaja Ca'Marcanda Promis			2.25	*(2.95)*	1.63	*(2.39)*	1.63	*(2.58)*
33	de Ladoucette Pouilly Fume	3.20	*(2.93)*	6.10	*(1.58)*	5.33	*(2.31)*	5.21	*(2.45)*
34	Snoqualmie Sauvignon Blanc	6.60	*(0.80)*	6.33	*(0.94)*	7.00		6.65	*(0.76)*
35	Chaddsford Merican			1.00	*(2.10)*	0.56	*(0.83)*	0.63	*(1.49)*
36	Turkey Flat Butchers Block			1.30	*(2.15)*	0.56	*(0.83)*	0.75	*(1.56)*
37	Elmo Pio Asti	3.50	*(3.50)*	3.44	*(2.79)*	1.89	*(2.73)*	2.82	*(3.01)*
38	San Telmo Merlot	5.60	*(2.80)*	5.11	*(2.28)*	6.89	*(0.31)*	5.91	*(2.10)*
39	Williams Selyem Chardonnay	6.80	*(0.40)*	6.78	*(0.63)*	6.89	*(0.31)*	6.83	*(0.48)*
40	The Foundry Syrah	6.00	*(1.26)*	5.56	*(2.06)*	6.89	*(0.31)*	6.17	*(1.55)*
		4.12	*(3.22)*	4.23	*(2.92)*	4.50	*(3.10)*	4.31	*(3.06)*

	Name	Sparkling Connoisseur		Retailer		Winemaker		Total	
1	Louis Jadot Pommard			0.60	*(1.80)*			0.25	*(1.20)*
2	Nobilo "Icon" Sauvignon Blanc			0.60	*(1.80)*			0.25	*(1.20)*
3	Bonterra Chardonnay	0.40	*(0.80)*	0.60	*(1.80)*			0.33	*(1.25)*
4	Blue Mountain Icewine			0.10	*(0.30)*	0.11	*(0.31)*	0.08	*(0.28)*
5	Chateau Lafite Rothschild								
6	Darting Riesling Kabinett					0.33	*(0.67)*	0.13	*(0.44)*
7	Frog's Leap Rutherford			0.10	*(0.30)*			0.04	*(0.20)*
8	Rizzi Barbaresco			0.10	*(0.30)*			0.04	*(0.20)*
9	Blackstone Pinot Noir			0.60	*(1.80)*			0.25	*(1.20)*
10	Gnemiz Tocai Friulano			0.50	*(1.50)*	0.11	*(0.31)*	0.25	*(1.01)*
11	Beringer White Zinfandel					0.11	*(0.31)*	0.04	*(0.20)*
12	Mount Nittany Bergwein			0.30	*(0.90)*	0.22	*(0.42)*	0.21	*(0.64)*
13	Louis Jadot Pouilly Fuisse	0.40	*(0.80)*	0.60	*(1.80)*	0.22	*(0.63)*	0.42	*(1.29)*
14	Taylor Lake Country Red								
15	Villa Sandi Prosecco	6.20	*(0.75)*	5.70	*(1.27)*	5.78	*(1.31)*	5.83	*(1.21)*
16	Straccali Chianti								
17	Barefoot Chardonnay			0.60	*(1.80)*			0.25	*(1.20)*
18	Conde de Vimioso			0.10	*(0.30)*			0.04	*(0.20)*
19	Veuve Clicquot Ponsardin Brut	7.00		6.50	*(1.50)*	7.00		6.79	*(1.00)*
20	Gloria Ferrer Chard. (Estate)	0.40	*(0.80)*	0.60	*(1.80)*			0.33	*(1.25)*
21	Bremerton Tamblyn								
22	B.P. Rothschild Pinot Noir	0.40	*(0.80)*	0.60	*(1.80)*			0.33	*(1.25)*
23	Chateau Frank Brut	6.60	*(0.49)*	6.10	*(1.22)*	6.89	*(0.31)*	6.50	*(0.91)*
24	Leapfrogmilch					0.22	*(0.42)*	0.08	*(0.28)*
25	Gloria Ferrer Chard. (Reserve)	0.40	*(0.80)*	0.60	*(1.80)*			0.33	*(1.25)*
26	McWilliam's Shiraz			0.10	*(0.30)*			0.04	*(0.20)*
27	Blue Mountain Riesling					0.22	*(0.42)*	0.08	*(0.28)*
28	Vendange Merlot								
29	Twin Fin Cabernet Sauvignon			0.10	*(0.30)*			0.04	*(0.20)*
30	B.P. Rothschild Cab. Sauv.			0.10	*(0.30)*			0.04	*(0.20)*
31	3 Bridges Botrytis Semillon			0.30	*(0.90)*	0.11	*(0.31)*	0.17	*(0.62)*
32	Gaja Ca'Marcanda Promis								
33	de Ladoucette Pouilly Fume			0.60	*(1.80)*	0.22	*(0.63)*	0.33	*(1.25)*
34	Snoqualmie Sauvignon Blanc			0.60	*(1.80)*			0.25	*(1.20)*
35	Chaddsford Merican								
36	Turkey Flat Butchers Block								
37	Elmo Pio Asti	4.80	*(2.86)*	5.70	*(1.27)*	5.89	*(1.37)*	5.58	*(1.80)*
38	San Telmo Merlot								
39	Williams Selyem Chardonnay	0.40	*(0.80)*	0.60	*(1.80)*			0.33	*(1.25)*
40	The Foundry Syrah			0.50	*(1.50)*			0.21	*(1.00)*
		0.68	*(1.93)*	0.84	*(2.09)*	0.69	*(1.95)*	0.75	*(2.01)*

	Name	Sweet						Total	
		Connoisseur		Retailer		Winemaker			
1	Louis Jadot Pommard	0.25	*(0.43)*	0.29	*(0.70)*			0.15	*(0.48)*
2	Nobilo "Icon" Sauvignon Blanc	0.25	*(0.43)*	0.57	*(1.05)*	0.44	*(0.68)*	0.45	*(0.80)*
3	Bonterra Chardonnay	0.50	*(0.87)*	0.57	*(1.40)*	0.33	*(0.67)*	0.45	*(1.02)*
4	Blue Mountain Icewine	5.25	*(3.03)*	6.86	*(0.35)*	7.00		6.60	*(1.53)*
5	Chateau Lafite Rothschild	0.25	*(0.43)*					0.05	*(0.22)*
6	Darting Riesling Kabinett	5.00	*(2.92)*	6.43	*(0.73)*	2.00	*(2.05)*	4.15	*(2.80)*
7	Frog's Leap Rutherford	0.25	*(0.43)*					0.05	*(0.22)*
8	Rizzi Barbaresco	0.50	*(0.87)*			0.11	*(0.31)*	0.15	*(0.48)*
9	Blackstone Pinot Noir	0.25	*(0.43)*	0.14	*(0.35)*			0.10	*(0.30)*
10	Gnemiz Tocai Friulano			0.86	*(2.10)*	0.57	*(0.90)*	0.59	*(1.50)*
11	Beringer White Zinfandel	2.50	*(2.87)*	4.43	*(2.32)*	4.44	*(1.95)*	4.05	*(2.42)*
12	Mount Nittany Bergwein	1.50	*(1.50)*	4.29	*(1.28)*	3.78	*(2.30)*	3.50	*(2.11)*
13	Louis Jadot Pouilly Fuisse	0.25	*(0.43)*	0.43	*(1.05)*	0.11	*(0.31)*	0.25	*(0.70)*
14	Taylor Lake Country Red	0.25	*(0.43)*	3.14	*(2.17)*	3.67	*(2.54)*	2.80	*(2.50)*
15	Villa Sandi Prosecco	0.50	*(0.87)*	2.86	*(2.17)*	3.00	*(2.55)*	2.42	*(2.37)*
16	Straccali Chianti	0.25	*(0.43)*					0.05	*(0.22)*
17	Barefoot Chardonnay	0.50	*(0.87)*	0.43	*(0.73)*	0.33	*(0.67)*	0.40	*(0.73)*
18	Conde de Vimioso	0.25	*(0.43)*	0.29	*(0.70)*	0.25	*(0.66)*	0.26	*(0.64)*
19	Veuve Clicquot Ponsardin Brut	0.50	*(0.87)*	1.14	*(1.88)*	0.33	*(0.67)*	0.65	*(1.31)*
20	Gloria Ferrer Chard. (Estate)	0.50	*(0.87)*	0.43	*(1.05)*	0.22	*(0.42)*	0.35	*(0.79)*
21	Bremerton Tamblyn	0.25	*(0.43)*	0.14	*(0.35)*	0.11	*(0.31)*	0.15	*(0.36)*
22	B.P. Rothschild Pinot Noir	0.25	*(0.43)*	0.14	*(0.35)*			0.10	*(0.30)*
23	Chateau Frank Brut	0.25	*(0.43)*	0.86	*(1.73)*	0.33	*(0.67)*	0.50	*(1.16)*
24	Leapfrogmilch	3.75	*(2.49)*	4.71	*(2.05)*	3.88	*(2.52)*	4.16	*(2.39)*
25	Gloria Ferrer Chard. (Reserve)	0.50	*(0.87)*	0.29	*(0.70)*	0.11	*(0.31)*	0.25	*(0.62)*
26	McWilliam's Shiraz	0.25	*(0.43)*	0.14	*(0.35)*	0.11	*(0.31)*	0.15	*(0.36)*
27	Blue Mountain Riesling	2.00	*(2.00)*	4.43	*(2.13)*	3.67	*(2.31)*	3.60	*(2.35)*
28	Vendange Merlot	0.25	*(0.43)*	0.14	*(0.35)*			0.10	*(0.30)*
29	Twin Fin Cabernet Sauvignon	0.25	*(0.43)*			0.11	*(0.31)*	0.10	*(0.30)*
30	B.P. Rothschild Cab. Sauv.	0.25	*(0.43)*					0.05	*(0.22)*
31	3 Bridges Botrytis Semillon	6.75	*(0.43)*	7.00		7.00		6.95	*(0.22)*
32	Gaja Ca'Marcanda Promis	0.50	*(0.87)*			0.14	*(0.35)*	0.17	*(0.50)*
33	de Ladoucette Pouilly Fume	0.50	*(0.87)*	0.43	*(1.05)*	0.11	*(0.31)*	0.30	*(0.78)*
34	Snoqualmie Sauvignon Blanc	0.25	*(0.43)*	0.57	*(1.05)*	0.44	*(0.68)*	0.45	*(0.80)*
35	Chaddsford Merican	0.50	*(0.87)*	0.71	*(1.75)*			0.35	*(1.15)*
36	Turkey Flat Butchers Block	0.25	*(0.43)*	0.14	*(0.35)*	0.11	*(0.31)*	0.15	*(0.36)*
37	Elmo Pio Asti	4.25	*(2.68)*	5.14	*(1.55)*	5.63	*(1.32)*	5.16	*(1.84)*
38	San Telmo Merlot	0.25	*(0.43)*					0.05	*(0.22)*
39	Williams Selyem Chardonnay	0.50	*(0.87)*	0.14	*(0.35)*	0.22	*(0.42)*	0.25	*(0.54)*
40	The Foundry Syrah	0.25	*(0.43)*			0.11	*(0.31)*	0.10	*(0.30)*
		1.04	*(2.01)*	1.45	*(2.42)*	1.20	*(2.26)*	1.26	*(2.28)*

	Name	Wines								Grand Total	
		Connoisseur		Retailer		Winemaker		Total			
1	Louis Jadot Pommard	7.00		6.20	*(1.25)*	7.00		6.67	*(0.90)*	2.25	*(3.01)*
2	Nobilo "Icon" Sauvignon Blanc	7.00		5.40	*(1.69)*	7.00		6.33	*(1.34)*	2.78	*(3.02)*
3	Bonterra Chardonnay	6.80	*(0.40)*	5.90	*(1.30)*	6.67	*(0.67)*	6.38	*(1.03)*	2.00	*(2.71)*
4	Blue Mountain Icewine	6.80	*(0.40)*	5.22	*(1.75)*	6.75	*(0.43)*	6.14	*(1.39)*	2.15	*(2.87)*
5	Chateau Lafite Rothschild	7.00		6.40	*(1.02)*	6.89	*(0.31)*	6.71	*(0.73)*	2.54	*(3.24)*
6	Darting Riesling Kabinett	7.00		5.50	*(1.69)*	7.00		6.35	*(1.34)*	2.83	*(3.06)*
7	Frog's Leap Rutherford	7.00		6.10	*(1.22)*	7.00		6.63	*(0.90)*	1.77	*(2.72)*
8	Rizzi Barbaresco	7.00		6.00	*(1.34)*	6.89	*(0.31)*	6.54	*(1.00)*	2.68	*(3.14)*
9	Blackstone Pinot Noir	7.00		5.70	*(1.49)*	6.67	*(0.67)*	6.33	*(1.18)*	1.68	*(2.66)*
10	Gnemiz Tocai Friulano	6.60	*(0.80)*	5.89	*(1.59)*	6.75	*(0.43)*	6.36	*(1.19)*	2.96	*(3.10)*
11	Beringer White Zinfandel	6.60	*(0.80)*	5.50	*(1.63)*	6.44	*(1.26)*	6.08	*(1.44)*	1.99	*(2.83)*
12	Mount Nittany Bergwein	6.60	*(0.80)*	4.89	*(1.37)*	6.50	*(1.32)*	5.86	*(1.49)*	1.68	*(2.53)*
13	Louis Jadot Pouilly Fuisse	7.00		6.00	*(1.34)*	7.00		6.58	*(1.00)*	2.70	*(2.98)*
14	Taylor Lake Country Red	6.20	*(0.98)*	4.67	*(1.76)*	6.33	*(1.56)*	5.65	*(1.73)*	1.21	*(2.32)*
15	Villa Sandi Prosecco	6.80	*(0.40)*	5.89	*(1.29)*	6.75	*(0.43)*	6.41	*(0.98)*	3.69	*(2.90)*
16	Straccali Chianti	6.80	*(0.40)*	5.80	*(1.47)*	6.89	*(0.31)*	6.42	*(1.11)*	2.48	*(3.05)*
17	Barefoot Chardonnay	7.00		6.20	*(1.25)*	6.56	*(0.96)*	6.50	*(1.04)*	2.10	*(2.82)*
18	Conde de Vimioso	6.80	*(0.40)*	6.00	*(1.32)*	6.63	*(0.70)*	6.43	*(1.00)*	1.59	*(2.60)*
19	Veuve Clicquot Ponsardin Brut	6.80	*(0.40)*	6.00	*(1.26)*	6.89	*(0.31)*	6.50	*(0.96)*	3.61	*(3.19)*
20	Gloria Ferrer Chard. (Estate)	7.00		6.00	*(1.18)*	7.00		6.58	*(0.91)*	2.03	*(2.74)*
21	Bremerton Tamblyn	7.00		6.22	*(1.23)*	6.78	*(0.42)*	6.61	*(0.87)*	1.30	*(2.32)*
22	B.P. Rothschild Pinot Noir	6.60	*(0.80)*	5.60	*(1.56)*	6.56	*(0.96)*	6.17	*(1.31)*	2.08	*(2.81)*
23	Chateau Frank Brut	6.60	*(0.49)*	5.90	*(1.37)*	6.89	*(0.31)*	6.42	*(1.04)*	2.41	*(2.91)*
24	Leapfrogmilch	6.80	*(0.40)*	5.40	*(1.50)*	6.44	*(1.26)*	6.08	*(1.38)*	1.83	*(2.60)*
25	Gloria Ferrer Chard. (Reserve)	7.00		5.80	*(1.47)*	7.00		6.50	*(1.12)*	2.00	*(2.73)*
26	McWilliam's Shiraz	7.00		6.20	*(1.17)*	6.89	*(0.31)*	6.63	*(0.86)*	1.77	*(2.72)*
27	Blue Mountain Riesling	6.80	*(0.40)*	5.40	*(1.91)*	6.75	*(0.43)*	6.17	*(1.46)*	2.29	*(2.85)*
28	Vendange Merlot	7.00		6.00	*(1.55)*	6.56	*(0.96)*	6.42	*(1.22)*	1.66	*(2.69)*
29	Twin Fin Cabernet Sauvignon	7.00		5.56	*(1.42)*	6.44	*(0.96)*	6.22	*(1.21)*	1.66	*(2.63)*
30	B.P. Rothschild Cab. Sauv.	6.60	*(0.80)*	6.00	*(1.55)*	6.56	*(0.96)*	6.33	*(1.25)*	2.27	*(2.94)*
31	3 Bridges Botrytis Semillon	6.80	*(0.40)*	5.50	*(1.75)*	6.89	*(0.31)*	6.29	*(1.34)*	2.39	*(2.98)*
32	Gaja Ca'Marcanda Promis	7.00		6.00	*(1.18)*	6.78	*(0.42)*	6.50	*(0.91)*	2.39	*(3.05)*
33	de Ladoucette Pouilly Fume	7.00		5.90	*(1.22)*	7.00		6.54	*(0.96)*	3.15	*(3.08)*
34	Snoqualmie Sauvignon Blanc	7.00		5.50	*(1.57)*	6.78	*(0.63)*	6.29	*(1.27)*	2.77	*(3.02)*
35	Chaddsford Merican	6.60	*(0.80)*	5.44	*(1.64)*	6.78	*(0.42)*	6.22	*(1.28)*	1.24	*(2.33)*
36	Turkey Flat Butchers Block	6.80	*(0.40)*	6.33	*(1.25)*	6.89	*(0.31)*	6.65	*(0.87)*	1.22	*(2.28)*
37	Elmo Pio Asti	6.60	*(0.49)*	5.80	*(1.33)*	6.89	*(0.31)*	6.38	*(1.03)*	3.76	*(2.96)*
38	San Telmo Merlot	7.00		6.10	*(1.45)*	6.88	*(0.33)*	6.57	*(1.06)*	1.64	*(2.65)*
39	Williams Selyem Chardonnay	7.00		6.40	*(0.92)*	7.00		6.75	*(0.66)*	2.10	*(2.84)*
40	The Foundry Syrah	7.00		6.40	*(1.20)*	6.78	*(0.42)*	6.67	*(0.85)*	1.60	*(2.62)*
		6.85	*(0.46)*	5.82	*(1.48)*	6.78	*(0.66)*	6.40	*(1.16)*	2.21	*(2.89)*

APPENDIX U:

ANALYSES OF NUMBER OF EXTENSIONS FOR INFERENCE TASK

A three-way ANOVA (Expert Group x Property x Premise Item) revealed a main effect for property $F(4,420) = 4.19$, $p = .002$, qualified by an interaction between property and group, $F(8,420) = 4.508$, $p = .00$. Tukey HSD post-hoc comparisons establish that the effect of property is due to a difference in the number of extensions (presented here as percent of group to correct for differences in the number of experts in each group) for fungus relative to dish, aliens, and Q. The difference between fungus and X was not significant.

Table U1 . Post-hoc tests of count means by property

Comparisons of Percent Count

Tukey HSD

(I) Property	(J) Property	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Fungus	Aliens	.11*	.03	.014	.015	.197
	Dish	.11*	.03	.014	.014	.196
	Q	.10*	.03	.034	.005	.187
	X	.08	.03	.102	-.009	.173

Based on observed means.
The error term is Mean Square(Error) = .053.
* The mean difference is significant at the 0.05 level.

As for the interaction, the graph below shows some dramatic differences for certain groups on certain properties. Winemakers and connoisseurs have a relatively high proportion of extensions for fungus, winemakers also do for Q, and retailers do for X.

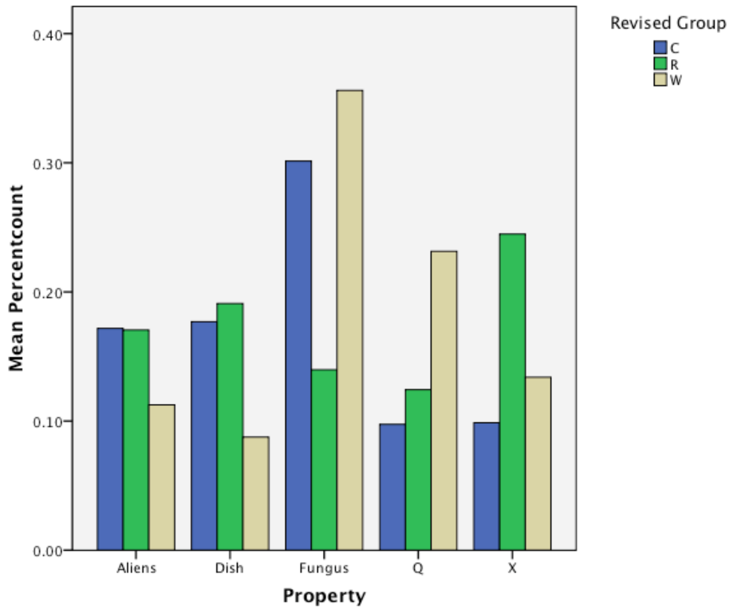


Figure U1. Average proportion (of other 39 wines) chosen by group for each property trial
Note: Groups are connoisseurs (C), retailers (R), and winemakers (W).

However, there are some problems with using the number (or proportion) of extension as a dependent variable. Differences among items are likely dependent at least partly on the wine set. People could be thinking the same way about 25 and 1, but show very different quantitative patterns because there are more chardonnays than pinots in the set.[47] Also, some of the apparent patterns may be the effect of a few individuals using strategies that result in a streak of "all" choices, extending the property to all wines in the set.

As it turns out, the peaks in the graph (of percent extensions) correspond to the conditions where people most often extended to all 39 other wines.

[47] This issue would probably also hold for other researchers who have used this approach (e.g., Shafto & Coley, 2003).

Table U2. Proportion of respondents who extended a property to "all" (39) other wines

	Aliens	Dish	Fungus	Q	X
Connoisseurs	0	0	0.200	0	0
Retailers	0.025	0.025	0	0	0.150
Winemakers	0	0	0.167	0.139	0

Table U2 shows the percentage of trials (for each group) that yielded an “all other wines” response. (Note: this means all other wines in the set, not necessarily all other wines in the world.) These items (100% of the 39 wines) may be swamping more subtle variation. Four individuals account for most of these cases (19 of 23 “all” responses). Each of these experts used a particular strategy consistently across each of the four premises paired for a given property. Interestingly, none of them used this strategy for more than one property: one used it for Q, one for X and two for fungus. Some of them also used it another isolated time (or two) for other properties. So, for example, in response to the fungus property, two individuals (a winemaker and a connoisseur) reasoned that one should test all wines because a fungus could be devastating to the industry (this had happened in the past). One winemaker reasoned that property “Q” must have something to do with alcohol, which is obviously present in all of the wines. For the chemical compound xergia, a retailer reasoned that it must have something to do with the winemaking process. Note that while experts extended properties to no other wines more frequently than to all wines (59 vs. 23 instances), no expert consistently responded with “none” for a given property type; there were no “streaks.” Thus, the quantity of choices may be less meaningful than the pattern of those choices.

APPENDIX V:

SORT DISTANCE TO INFERENCE CHOICE CORRELATIONS BY EXPERT GROUP

Table V1. Mean correlation of premise-target distance (first sort) with target choice

	Connoisseurs			Retailers			Winemakers		
	M	*SD*	*t (df)*	*M*	*SD*	*t (df)*	*M*	*SD*	*t (df)*
Aliens	$-.28^{**}$	.11	-5.91 (4)	$-.16^{**}$	.13	-4.13 (9)	$-.33^{**}$	.25	-3.76 (7)
Dish	$-.25^{**}$	.08	-7.27 (4)	$-.14^{**}$	.08	-5.87 (9)	$-.26^{**}$	.21	-3.72 (8)
Fungus	$-.47^{*}$	.24	-4.00 (3)	$-.19^{*}$	.21	-2.89 (9)	-.15	.30	-1.50 (8)
Q	$-.38^{**}$	*.15*	-5.46 (4)	$-.22^{**}$	.16	-4.31 (9)	$-.20^{*}$	.17	-3.15 (6)
Xergia	$-.34^{**}$	.14	-5.54 (4)	$-.15^{*}$	.14	-3.18 (8)	$-.21^{**}$	.14	-4.73 (8)

$^{**}p < .01$, one-tailed
$^{*}p < .05$, one-tailed

As predicted, the average correlations were negative; except for one, they were statistically significant. The one non-significant correlation (winemakers-fungus) also had the largest standard deviation. Investigation of the individual winemakers' correlations for fungus revealed that the single highest (positive) correlation came from this set (.41). This value was a clear outlier; the next-highest correlation was .14. With the outlier removed, the result for winemakers-fungus was M = -.30, SD = .22, t(7) = -3.84, p < .01. A two-way, mixed ANOVA (Expert Group x Property) found no significant effects.

APPENDIX W:

MATCH SCORES FOR PREMISES

There was also a main effect of premise (Figure X). All pair-wise comparisons were significant (*p*s < .05 using Bonferroni adjustments). Matches (across all types) were most influential for premise wine #19, followed by #1, then #25, and least so for #14.

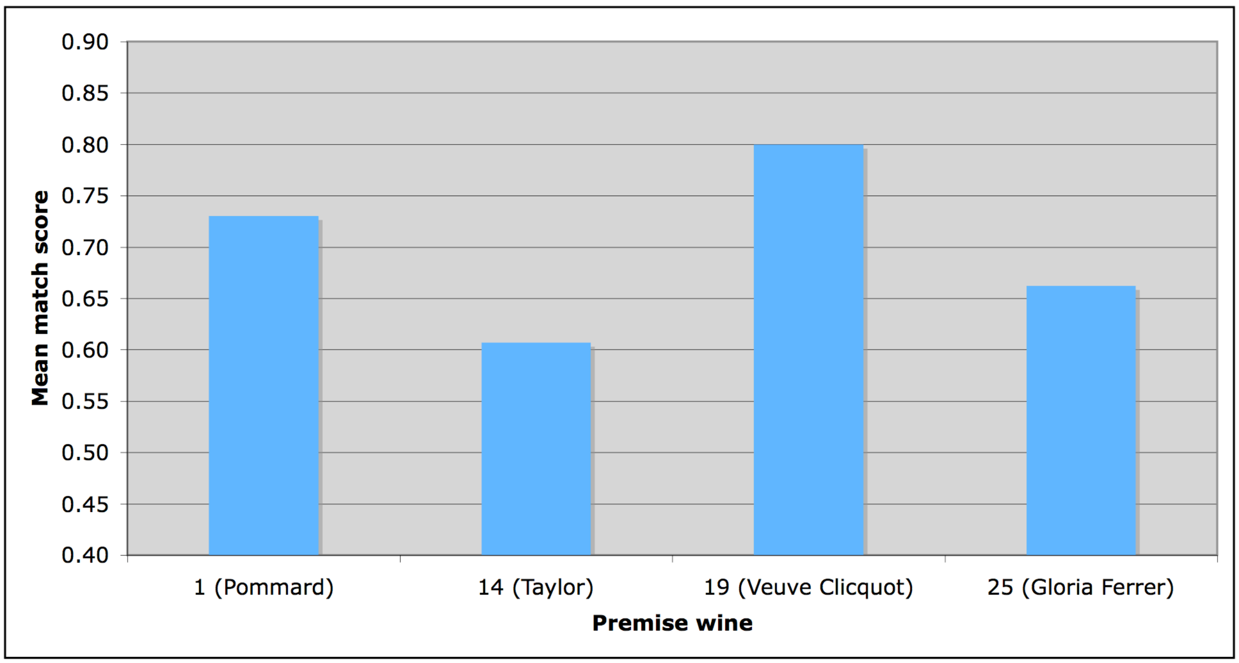


Figure W1. Mean score presented by premise wine

Note: This chart presents mean scores, averaged across all participants, properties, and match types. The scores are the proportion of selections (of the 39 choices for each trial) that represented a match-consistent (L2) response—either the selection of a wine that matches the premise or the rejection of a wine that does not match the premise.

APPENDIX X:

JUSTIFICATION CODES BY PROPERTY

Table X1. Percent of group justifications as code type by property

Code Type	Property					Post-hoc tests	
	Aliens	Dish	Fungus	Q	Xergia	*F*(4,10)	Tukey *HSD*, *p*s < .05
color	0.295	0.362	0.063	0.161	0.119	20.31	AD>FXQ
grape	0.148	0.197	0.155	0.203	0.298	5.34	X>AF
other	0.238	0.236	0.251	0.205	0.070	ns	
price/quality	0.406	0.353	0.016	0.164	0.061	6.16	A>XF; D>F
process	0.065	0.067	0.186	0.295	0.387	5.78	X>AD
region	0.076	0.079	0.405	0.252	0.189	10.14	F>ADX
style	0.409	0.394	0.016	0.109	0.072	32.38	AD>FXQ
type	0.305	0.247	0.067	0.220	0.160	ns	