

OVERHEAD COST POOL CLASSIFICATION AND JUDGMENT PERFORMANCE

M. Christian Mastilak
Department of Accountancy
Xavier University
3800 Victory Parkway
Cincinnati, OH 45207
mastilakc@xavier.edu

September 4, 2007

This paper is based on my dissertation at Michigan State University. I gratefully acknowledge the help of my committee, Joan Luft (chair), Rick DeShon (Department of Psychology), Mike Shields and Cheri Speier. I am also thankful for the help of Sue Haka, Fred Jacobs, Ranjani Krishnan, Tom Linsmeier, Kathy Petroni, Fabienne Miller, and Alex Woods on this and earlier versions of this paper, as well as comments from workshop participants at Case Western Reserve University, George Mason University, Miami University, Michigan State University, Wayne State University, Western Illinois University, and the University of Wisconsin-Milwaukee. Financial support from the Michigan State University Department of Accounting and Information Systems, the Michigan State University Graduate School, PricewaterhouseCoopers, and Ernst & Young is gratefully acknowledged.

OVERHEAD COST POOL CLASSIFICATION AND JUDGMENT PERFORMANCE

Abstract

This paper addresses the research question: does an organization's choice of overhead cost pool classification affect individuals' predictive judgments about overhead costs? I use psychology research to develop hypotheses about how overhead cost pool classification affects the expected accuracy of participants' predictions of a target overhead cost based on their estimation of the coefficients relating the target overhead cost to four predictor overhead costs. The predictor costs are located either within the same overhead pool as the target cost or in another pool. I vary relationships among the costs such that the larger predictor-target coefficient is either within a single cost pool or across different cost pools, and the relationships are positive or negative. I hypothesize that implicit estimates of coefficients on predictors within the same pool as the target cost will be more accurate than those on predictors in a different pool. I also hypothesize that predictions of the target cost are affected by an interaction of the location of the stronger predictor and the sign of the relationships. As hypothesized, participants implicitly estimate coefficients within a pool more accurately than across pools. Further, the location of the stronger predictor and the sign of the relations interact, such that the effect of the location of the stronger predictor on judgment accuracy is greater when relationships among costs are negative than when they are positive. I thus provide theory-consistent evidence of a cognitive effect of overhead cost pool classification in addition to informational effects modeled by the analytic literature.

Keywords: cost pools, classification, judgment, cost estimation

Data availability: Contact the author

I. INTRODUCTION

A fundamental function of accounting is the combination of individual accounts into groups of accounts. For example, external financial reporting and tax return preparation involve the reporting of hundreds or thousands of individual accounts in several dozen lines on financial statements or tax returns. Similarly, organizations classify individual overhead cost accounts into overhead cost pools. In this paper I examine how different overhead cost pool classification decisions affect judgment performance of individuals who use accounting reports. Throughout this paper, I use the label “overhead cost pool classification” to describe the assignment of individual overhead cost accounts to specific overhead cost pools.

Managers make judgments about costs based in part on information about other costs (Datar et al. 1993; Joshi et al. 2001). The accuracy of these judgments may be affected by organizations’ overhead cost pool classification decisions. For example, Joshi et al. (2001) found that managers of steel firms grossly underestimated their environmental regulation cost burden. Part of the cost burden was “hidden” in non-regulatory cost pools, such as material or labor cost pools. Managers at these firms were aware that a portion of the burden was “hidden,” but their estimates of the hidden portion were inaccurate. The managers suggested that performance evaluation and plant closure decisions were likely affected by these inaccurate estimates of regulatory costs. Whether the firms’ overhead cost pool classification decisions were a *cause* or an *effect* of managers’ understanding of their firms’ costs is not addressed in Joshi et al. (2001). Further, to my knowledge, there is no research that addresses the issue of whether overhead cost pool classification decisions affect managers’ judgment performance. Thus, this paper addresses that question.

In this paper I examine a setting in which individuals receive reports about both overhead cost pool totals and individual overhead cost accounts within the pools. Individuals use these reports in a budgeting task that requires the prediction of an overhead cost (the “target cost”). Consistent with empirical evidence on managerial decisions (e.g., Datar et al. 1993), I focus on judgments about relationships among individual overhead costs within and across overhead pools, rather than judgments about relationships between overhead costs and cost allocation bases (i.e., cost drivers). I use psychology research to develop hypotheses about how overhead cost pool classification affects the accuracy of participants’ predictions of a target overhead cost based on other overhead costs that may be predictive of the target (the “predictors”).¹

I use an experimental research design in which participants make predictions about the target based on information about the predictors. The predictor costs are located either within the same overhead pool as the target cost or in another pool. Participants’ predictions of the target cost are based on their implicit estimates of the coefficients relating the target to the various predictors.² I vary the relationships among costs in two ways. First, the location of the stronger predictor-target relationship is varied such that it is either within a single cost pool or across different cost pools. Second, the relationships among costs are either positive or negative. These manipulations allow me to examine the effects on judgment of different overhead cost pool

¹ For example, in the setting examined by Joshi et al. (2001), the total cost burden of environmental regulations would be the “target” cost, and information about the other cost pools (materials, labor, etc.) would be the “predictors.” Managers would use information about the other cost pools to predict the total regulatory cost burden.

² I model participants’ judgments using regression and generate estimates of the intercept and implicit coefficients in their prediction models. Participants do not directly estimate the coefficients relating the target cost to the predictors.

classification decisions in settings with differing cost behaviors (e.g., costs that are substitutes vs. complements).³

Based on psychology research (e.g., Brehmer 1974; Broniarczyk and Alba 1994; Sniezek 1986), I predict that overhead cost pool classification affects judgment performance by directing individuals' attention toward within-pool relationships and away from across-pool relationships. Thus, I hypothesize that individuals will implicitly estimate within-pool coefficients more accurately than across-pool coefficients. I further expect that relationships to which less attention is allocated are *underestimated* rather than randomly misestimated (e.g., Broniarczyk and Alba 1994). Thus, individuals will tend to underestimate across-pool predictors more than within-pool predictors. I also predict that accuracy of individuals' predictions of the target overhead cost is an interactive function of the sign of the relationships among costs and the location of the stronger predictor, such that the effect of the location of the stronger predictor on judgment accuracy is greater when relationships among costs are negative than when they are positive.

The experimental results are generally consistent with hypotheses. First, participants' implicit estimates of the coefficients relating predictor and target overhead costs are more accurate when predictor and target are within the same pool than when they are across pools. Second, participants underweight across-pool predictors more than within-pool predictors. Further, the sign and strength of the relationships among predictors and the target interact to affect the accuracy of participants' predictions of the target. The effect of overhead cost pool classification on judgment accuracy is greater when relationships among costs are negative than when they are positive. Finally, supplemental analysis provides evidence that overhead cost pool

³ In this paper, I use the term "cost behavior" to describe whether the variable portion of target and predictor overhead costs are positively or negatively correlated with each other and whether the relation between target and predictor is weak or strong.

classification affects the allocation of attention across different predictors, and the allocation of attention affects prediction accuracy.

This study considers subjective judgments about overhead costs. Subjective judgment is used often in organizations (MacKinnon and Bruns 1992; Kaplan and Norton 1996; Ittner et al. 2003), and thus it is an important topic for research. Even a formal statistical analysis such as a multiple regression analysis will only increase judgment accuracy if the appropriate relationships are examined (i.e., if appropriate variables and functional form are used in the analysis). When performing formal statistical analyses, individuals often use subjective judgment to choose relationships to examine and the functional form of those relationships. My results imply that overhead cost pool classification can affect *which* relationships are examined as well as *how accurately* they are estimated. Thus, overhead cost pool classification may influence formal statistical analyses of overhead costs as well as subjective judgments.

The rest of this paper is organized as follows. Section 2 presents the literature review and hypothesis development. Section 3 describes the research design. Section 4 presents results, and section 5 concludes.

II. LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

First, I introduce accounting literature on overhead cost pool classification. Then, I will provide an example of a judgment task that is affected by organizations' overhead cost pool classification decisions. Next, I explain why these overhead cost pool classification decisions vary across organizations. Then, I use findings from psychology literature to describe (1) how variation in overhead cost pool classification will affect individuals' beliefs about costs, (2) how those beliefs will affect individuals' allocation of their attention when they perform the task, and

(3) how attention allocation affects judgment performance in the task. Further, I make predictions about how different cost behaviors will affect judgment performance in the task.

Overhead cost pool classification

Organizations' decisions regarding the allocation of overhead costs to cost objects can be modeled as a two-stage process, as shown in Figure 1 (Cooper and Kaplan 1999).

Insert Figure 1 here.

Prior academic literature on overhead cost pool classification has focused primarily on the second stage. Analytical literature has considered the effect of interactions between organizations' cost functions and overhead cost pool classification on the cost allocation error reported by various cost allocation systems (e.g., Noreen 1991; Datar and Gupta 1994; Christensen and Demski 1995, 1997; Bromwich and Hong 1999). Empirical literature has typically examined the validity of claims that important cost drivers other than production volume exist (e.g., Foster and Gupta 1990; Banker et al. 1995; Anderson 1995; Ittner et al. 1997). Experimental literature has examined the effects of cost system accuracy, economic complexity and different types of feedback on cost forecasting and product pricing/output decisions (Gupta and King 1997; Briers et al. 1999), and the interaction of cost system and incentives on innovation, efficiency and profitability in teams (Drake et al. 1999). To my knowledge, there is no literature on the effects of overhead cost pool classification on individuals' judgment performance about individual overhead costs.

The potential for an effect of overhead cost pool classification on individuals' judgment performance is an important topic for research. There is evidence that overhead costs can be positively or negatively related to costs in their own and other overhead cost pools in statistically

and economically significant ways. For example, Anderson (1995) documents strong negative correlations among setup costs, accounted for in a setup cost pool, and power costs, accounted for in an administrative cost pool. Ittner et al. (1997) document significant positive and negative correlations among costs in different pools and at different levels of the activity-based costing hierarchy. And Joshi et al. (2001) report that for a sample of U.S. steel firms, large portions of environmental regulation compliance costs are “hidden” in non-regulatory cost pools.

Variations in overhead cost pool classification

Relationships among overhead costs within and across cost pools will vary across organizations because organizations have different purposes for classifying overhead costs into pools. Two possible purposes are discussed here: product costing and responsibility accounting.⁴

Product costing

One important purpose for classifying overhead costs into pools is the allocation of costs to products or services for such purposes as pricing, deciding whether to drop a product from the product line, determining optimal input mixes, and designing products to minimize cost (hereafter “product costing”). If product costing is the primary purpose of an organization’s cost accounting system, then the organization will likely choose to assign to the same cost pools those overhead costs that are strongly positively correlated with each other (Garrison et al. 2006; Horngren et al. 2002; Kaplan and Cooper 1996). Overhead cost pool classification decisions that result in pools containing costs with stronger positive within-pool correlations will minimize product costing errors (Noreen 1991; Christensen and Demski 1995; Bromwich and Hong 1999).

Noreen (1991, 163), discussing the requirements for a costing system to provide costs that are appropriate for product-drop and product-design decisions, writes,

⁴ Other reasons exist, such as inventory valuation for external reporting purposes, that may also affect overhead cost pool classification decisions (Kaplan and Cooper 1999).

[A] well-specified ABC system exists in which product costs are avoidable costs and activity costs are incremental costs if and only if: 1) the underlying cost function...can be partitioned into cost pools, each of which depends only upon a single activity; 2) the cost in each pool is strictly proportional to its activity, and 3) each activity can be divided among products in such a way that the portion attributed to each product depends only upon that product.”

If costs within a pool are all proportional to a single activity, then it follows that the costs will be highly correlated with each other. Bromwich and Hong (1999) and Christensen and Demski (1995) offer similar discussions of the requirements of pooled cost accounting system.

In summary, if product costing is the primary determinant of overhead cost pool classification decisions, it is likely that costs within pools will be strongly correlated, and costs within a pool will be weakly correlated or uncorrelated with those in other pools.

Responsibility accounting

Another important purpose for classifying overhead costs into pools is responsibility accounting (that is, a system under which costs for which an individual or responsibility center is responsible are pooled to provide a summary statistic for that individual's or responsibility center's performance). If responsibility accounting is a primary purpose for the cost system, within-pool relationships can be strong or weak. Responsibility accounting purposes for cost system design may result in weakly correlated overhead costs being classified into a single pool. This may occur when a single individual or responsibility center has responsibility for weakly correlated costs, and when measurement error creates difficulty in dividing those costs into different pools (e.g., Datar and Gupta 1994). For example, an employee may spend time on three different activities, and it is difficult to measure the time spent on each of two of those activities. The portions of that employee's compensation related to time spent on those two activities might then be recorded within a single pool even if they are weakly correlated, to prevent negative consequences that might result from allocating the time more finely but with error.

On the other hand, overhead costs in one pool may be strongly correlated with overhead costs in a different pool, because different individuals or responsibility centers are responsible for the two costs, and accounting for them in separate pools facilitates calculations such as variances and responsibility center-level reports (Garrison et al. 2006; Horngren et al. 2002; Cooper and Kaplan 1999). Thus, in an organization whose overhead cost pool classification decisions are based primarily on responsibility accounting, it is possible (but not necessary) that weakly correlated overhead costs will be classified into a single overhead cost pool and strongly correlated costs may be classified into different overhead cost pools.

Example of a prediction task affected by overhead cost pool classification

To illustrate the effects of overhead cost pool classification on judgment performance, consider the following example of a task an individual in an organization might face. As part of an organization's operating budgeting process, the individual must predict the next month's expected spoilage costs for each of several similar manufacturing plants. Unlike such costs as direct labor, spoilage costs are not the direct result of a spending decision, but rather an indirect result of a variety of other spending and operating decisions. The individual's task is to predict spoilage cost from levels of spending on other overhead costs based on relationships among overhead costs observed in, and learned from, a prior month's report. Specifically, the individual uses spending on inspection and supervision as predictors of spoilage costs. Assuming a linear, additive relationship between spoilage, inspection and supervision, the prediction can be modeled by $Spoilage = Constant + \beta_1 Inspection + \beta_2 Supervision + error$.

To successfully predict spoilage, the individual must estimate the constant (i.e., the fixed cost component of spoilage cost) and β_1 and β_2 , the coefficients on inspection and supervision (i.e., the components of spoilage cost that vary with inspection and supervision). The sign and

magnitudes of the coefficients will depend on the organization's cost function.⁵ That is, β_1 and β_2 may be positive or negative (or zero), and β_1 may be larger or smaller than β_2 .

The organization's overhead cost pool classification will affect whether inspection and/or supervision costs are in the same pool as spoilage costs. There are several overhead cost pool classification decisions the organization could make. Assume that inspection and supervision costs are uncorrelated and are in separate responsibility centers, so the organization has decided to account for these costs in separate pools. Further, assume the organization has decided to assign the three costs to two pools, and so must now decide whether to assign spoilage to the same pool as inspection pool or the same pool as supervision. Also assume that spoilage and inspection have a stronger relationship ($\beta_1 = 1.01$) and spoilage and supervision have a weaker relationship ($\beta_2 = 0.46$).⁶ Finally, assume that responsibility for and some influence over both spending on supervisors and reducing spoilage exist within a common responsibility center. Figure 2 presents the example graphically.

Insert Figure 2 here.

Given the assumptions and the discussion above, there are two possible overhead cost pool classification decisions the organization can make. If allocating spoilage, inspection and supervision costs to cost objects is of primary importance, then the organization would likely include spoilage in the same cost pool as inspection. Figure 2, Panel A depicts this overhead cost

⁵ The magnitude of the constant (the fixed portion of the overhead cost) will also depend on the cost function. However, in this paper I focus only on coefficients (the variable portion of the overhead cost).

⁶ For simplicity, differences between larger and smaller coefficients are assumed to be statistically significant. Further, costs are assumed to have similar scales, so larger coefficients imply stronger correlations and more predictive validity.

pool classification. Spoilage has a stronger relationship with inspection than does supervision; as noted above, assigning spoilage and inspection to the same pool will allow for a better choice of allocation base than will assigning spoilage and supervision to the same pool. Note that in this overhead cost pool classification, there is a *stronger* relationship *within* a cost pool (spoilage and inspection) and a *weaker* relationship *across* cost pools (spoilage and supervision).

Alternatively, if responsibility accounting is of primary importance, then the organization would likely assign spoilage and supervision to the same cost pool, to allow the cost system to support clear responsibility accounting. Figure 2, Panel B depicts this overhead cost pool classification. Note that in this overhead cost pool classification, there is a *weaker* relationship *within* a cost pool (spoilage and supervision) and a *stronger* relationship *across* cost pools (spoilage and inspection).

In summary, there is likely variation in the overhead cost pool classification decisions made by organizations, based on different purposes for the overhead cost pool classifications. Whether and how these overhead cost pool classification decisions matter depend on how overhead cost pool classification decisions affect individuals' judgment performance. The following subsection reviews the accounting and psychology literatures and develops hypotheses about how overhead cost pool classification decisions are expected to affect judgment performance.

Overhead cost pool classification and beliefs about costs

The classification of costs into cost pools results in individual cost accounts being grouped together and given a common label (such as "quality control costs"). The grouping of individual items under a label has been investigated by the psychology literature on categorization (e.g., Markman and Gentner 2001). Categorization has been found to be a

fundamental part of human information processing (Murphy and Medin 1985). The categories in which individual items reside, and the labels on those categories, have been found to influence the inferences individuals make about those individual items. Specifically, individuals infer that items given the same (a different) category label are similar (dissimilar) to other items given the same category label (Yamauchi and Markman 2000; Markman and Gentner 2001). Thus, when individuals make inferences about relationships among costs classified in the same (different) cost pools, they are likely to infer that the costs are similar (dissimilar) to each other.

Discussions of cost pool classification in accounting textbooks and literature typically focus on the correlations among costs within pools. Correlation among costs is the feature on which overhead cost pool classification is based (e.g., Christensen and Demski 1995; Kaplan and Cooper 1998; Horngren et al. 2002).⁷ For example, Horngren et al. (2002, p. 142) write:

Each of the activity-related cost pools [is] homogeneous. Why? Because each activity cost pool includes only a narrow and focused set of costs (for example, setup or distribution). Over time, the costs in each activity-cost pool have a cause-and-effect relationship with the cost-allocation base for that activity (for example, setup-hours in the case of setup costs and cubic feet of packages moved in the case of distribution costs).

Similarly, Kaplan and Cooper (1998, p. 86) cite an internal training manual saying, “A resource comprises a *distinct and homogeneous* grouping of existing costs fulfilling a similar function...” (emphasis added). Also, Roth and Borthick (1991, p. 39) write that the assumption of homogeneity in activity-based costing models “means that the costs in each pool are driven by a single activity or by highly correlated activities. *Highly correlated* means that changes in the level of one activity are accompanied by proportional changes in the other activities” (emphasis in original).

⁷ These examples refer to activity-based costing systems, and much of the discussion regarding overhead cost pool classification has been in the context of ABC systems. However, the choice of activities as the driver or allocation base is not required for the principle of overhead cost pool classification based on correlation to hold.

These passages strongly imply that because costs within each pool move with a single activity driver, they move together with each other. Further, because each pool's costs are "narrow and focused" or, alternatively, "distinct and homogeneous," costs that are not highly correlated are classified into separate pools. Thus, correlation among costs is an important feature of individual costs that individuals will likely infer is similar among costs classified within a cost pool, and dissimilar among costs classified into different cost pools. Thus, in the absence of strong evidence to the contrary, students and practitioners of accounting are likely to believe that overhead costs within a pool (across pools) are positively and relatively strongly correlated (not correlated).

Despite the assumption in the literature about strong relationships being classified within pools, organizations' overhead cost pool classification decisions vary, as described above. Therefore, individuals within organizations will sometimes use accounting information in which strong correlations are across-pool rather than within-pool, for example, when overhead cost pool classification is based on responsibility accounting. In these circumstances, individuals will have knowledge of some, but not all, of the relationships among costs. For those relationships for which individuals do not have current knowledge, individuals may rely on overhead cost pool classification as a source of beliefs about relationships among overhead costs.

The accuracy of individuals' estimates of relationships among overhead costs may be affected by organizations' overhead cost pool classification decisions and the relationships among overhead costs. Individuals' implicit estimations of relationships among overhead costs, and their resulting predictions of the target overhead cost, will be affected by both the allocation of attention among potential relationships and the cognitive difficulty of estimating those

relationships that receive attention. The following analysis describes how overhead cost pool classification and cost behavior can affect attention allocation and estimation accuracy.

Beliefs about costs, attention allocation and judgment performance

Beliefs about relationships in data can affect the allocation of individuals' attention among the potential relationships (Brehmer 1974; Broniarczyk and Alba 1994; Luft and Shields 2001). Research in accounting and psychology has documented the pervasive effects of individuals' beliefs on how they cognitively examine data and their resulting judgments in a variety of task settings, such as credit scoring, GPA and temperature prediction, consumer judgments of a price-quality relationship, and profit prediction (Muchinsky and Dudycha 1974, 1975; Sniezek 1986; Broniarczyk and Alba 1994; Luft and Shields 2001). Psychology research indicates that individuals examine potential relationships between a target and multiple predictors one predictor at a time, rather than simultaneously (Brehmer 1974). The order in which individuals examine relationships is influenced by their beliefs. Individuals examine potential relationships they believe to be more important earlier than relationships they believe to be less important (Brehmer 1974). Relationships examined earlier are estimated more accurately because individuals' limited cognitive capacity leads to fatigue, reduced cognitive effort and inconsistency (Brehmer 1974; Luft and Shields 2001). Thus, if individuals believe that overhead costs within (across) pools are likely (unlikely) to be related to each other, then overhead cost pool classification is expected to affect judgment accuracy. However, if individuals hold strong beliefs about relationships, but those relationships are not present in the data, those beliefs can overwhelm inspection of the data, individuals' judgments can be biased in the direction of their beliefs (Broniarczyk and Alba 1994).

Evidence from Joshi et al. (2001) is consistent with overhead cost pool classification affecting judgment performance among experienced managers. The authors found that large portions of the costs steel firms incur in complying with environmental regulations are accounted for in non-regulatory cost pools. These costs are misclassified because they are not easily identifiable as regulatory costs. For example, if a mill switches from a cheaper, high-polluting coal to a more-expensive, cleaner-burning coal, the entire cost of the coal is counted as a raw material cost, even though the difference in coal prices is due to regulation and would not be incurred save for the regulation. Managers at the firms studied are aware of the existence of across-pool relationships and the fact that some regulatory cost burden is “hidden.” Further, managers consider environmental regulation costs to be an important factor in their firms’ ability to compete. The managers thus have both the opportunity and the motivation to understand the magnitude of the relationships. However, the managers grossly underestimated the magnitude of regulatory costs – they estimated that about half the cost was hidden, while the authors estimate that up to 90% of the cost is hidden. In interviews, the managers said they believed their underestimation of the regulatory costs affected decisions such as performance evaluation and plant closure or continuance decisions. However, whether the overhead cost pool classification is the cause or the effect of managers’ beliefs about the magnitude of “hidden” regulatory costs is uncertain.⁸

Effect of cost behavior on judgment performance

The sign and strengths of relationships can affect the accuracy of individuals’ estimations of the relationships independently of attention-allocation effects (Brehmer 1974; Sniezek 1986).

⁸ That is, managers may have had beliefs that the amount of potentially hidden regulatory cost was small, and designed their regulatory cost pools based on those beliefs. The design of Joshi et al. (2001) does not support inferences about causality between overhead cost pool classification and managers’ beliefs.

Even when individuals have correct beliefs about the sign of relationships, positive relationships are estimated more accurately than are negative relationships, indicating an inherent difference in the cognitive difficulty of estimating positive vs. negative relationships, all else equal. Thus, when individuals attempt to revise beliefs about relationships among overhead costs, they are likely to do so more accurately when the costs are positively correlated than when they are negatively correlated. Further, psychology research indicates that relatively weak relationships (i.e., those with correlations below approximately 0.6) are generally difficult for individuals to estimate subjectively, while stronger relationships (i.e., those with correlations above approximately 0.6) are generally estimated more accurately (e.g., Broniarczyk and Alba 1994). Thus, the sign and strength of a relationship affect the accuracy with which it is estimated.

There is also evidence that beliefs about relationships and the sign of relationships interact to affect judgment performance. Sniezek (1986, experiment 3) found that the effect of incorrect beliefs on judgment performance was greater when the relationships in the data were negative than when they were positive.

Consistent with psychology and accounting research, I use a lens model to analyze how and how well individuals learn and predict relationships from data.⁹ A conceptual depiction of the lens model is shown in Figure 3. The predictors in the center are the predictors individuals can use to predict the target. The environmental or true model is depicted on the left side of the figure; this model is derived from the data (Y_e and X_i) that are presented to individuals and that they could use to estimate the environmental coefficients (β_{ei} , where ei refers to the environmental coefficient on predictor i). The subjective model on the right side of the figure represents the model of individuals' subjective predictions of the target (Y_s) based on levels of

⁹ For a thorough and recent treatment of the lens model in accounting research, see Luft and Shields (2001).

the predictors. Lens model studies regress individuals' subjective predictions of the target on the predictors and estimate each participant's "implicit coefficients" (β_{sij} , where sij refers to participant j 's implicit coefficient on predictor i).

Insert Figure 3 here.

Hypothesis development

To aid the exposition of the hypotheses, I first briefly describe the experimental task. I examine an operating budgeting task in which participants predict the amount of the next month's expected spoilage costs in each of 20 similar manufacturing plants. The prediction is based on relationships among overhead costs that participants learn by examining the prior month's cost reports. The amount of reported spoilage costs is affected by two predictors (the "non-zero predictors").¹⁰ Two additional predictors are not actually predictive of spoilage (the "zero predictors"). The five overhead costs are assigned to two pools. Each pool contains one cost that is predictive of spoilage (a non-zero predictor) and one cost that is not predictive of spoilage (a zero predictor).

To examine whether the effect of overhead cost pool classification varies with variable cost behavior, I use two manipulations of cost behavior. First, I vary the pool location of the stronger predictor – the predictor with the larger coefficient is either within the same pool as spoilage or in the other pool (that is, across pools). Second, I vary the sign of the nonzero coefficients on the predictors – they are either positive or negative. Figure 4 depicts the overhead

¹⁰ Throughout this paper, I use the label "predictors" to indicate the costs that participants could use to predict the target cost regardless of whether they are statistically significant predictors. The inclusion of zero-coefficient predictors in a setting in which participants are likely to expect positive correlations allows for an empirical test of this departure from participants' likely expectations, and allows for tests to determine the effect of overhead cost pool classification on participants' estimation of coefficients.

cost pool classification and cost behavior conceptually. Panel A presents the condition in which the stronger predictor relationship is within the same pool as spoilage; Panel B presents the condition in which the stronger predictor relationship is across pools.

Insert Figure 4 here.

Participants are told the expected signs but not the magnitudes of the relationships between spoilage and the predictors. To perform well on the task, participants must use the cost reports to estimate the relationships between spoilage and the other overhead costs.¹¹ They must then apply those estimates of the relationships to the next month's budget of the four overhead costs and predict spoilage based on budgeted levels of the four other costs. An example of the experimental materials is shown in Appendix A.

Hypotheses 1 and 2 make predictions about errors in participants' implicit coefficients on individual predictors. Hypothesis 3 predicts how variation in overhead cost pool classification and cost behavior affect the overall prediction accuracy of overhead costs.

Hypotheses 1 and 2: Coefficient estimation errors

As described earlier, individuals generally examine potential relationships sequentially, with relationships believed to be more important examined earlier and more accurately than those believed to be less important (Brehmer 1974; Luft and Shields 2001). Consistent with the accounting literature reviewed above, I expect stronger (weaker) within-pool and weaker (stronger) across-pool relationships to be consistent (inconsistent) with individuals' beliefs about the relationships among overhead costs assigned to cost pools when the individuals have no other

¹¹ Participants must also implicitly estimate the constant in addition to the coefficients on the predictors. However, large errors in estimating the constant are unlikely, as they would result in predicted levels of spoilage that are outside the range of those in the prior month's data.

strong sources of belief about the relationships among the costs. Because individuals likely believe that overhead costs classified within the same pool are strongly related and overhead costs classified across different pools are not strongly related, they will devote more attention to within-pool relationships and estimate them more accurately than across-pool relationships. I expect this effect even though in my research design, all participants examine statistically identical data sets, with detailed cost information available in all conditions, and my research design includes relatively few potential predictors and offers a strong and direct financial incentive for optimal performance.

Beliefs can affect judgment accuracy in at least two ways. First, beliefs can affect the *bias* of individuals' judgments (Sniezek 1986; Billman et al. 1992; Broniarczyk and Alba 1994; Luft and Shields 2001). Because, in the absence of specific knowledge to the contrary, individuals likely believe that within-pool (across-pool) relationships exist or are more important (do not exist or are less important), they will exhibit a pattern of error biases consistent with these beliefs. For the non-zero predictors' coefficients, individuals will pay more attention to within-pool coefficients than to across-pool coefficients. Thus, individuals will estimate within-pool coefficients with relatively little bias. Because individuals will pay less attention to the across-pool coefficients, they tend to treat the across-pool coefficients as closer to zero than the data indicate. That is, individuals will underestimate across-pool non-zero predictors' coefficients more than within-pool non-zero predictors' coefficients.

For the zero predictors' coefficients, individuals will likely believe that the within-pool coefficients are non-zero. Broniarczyk and Alba (1994) present evidence that in the presence of strong beliefs about a relationship, individuals will make judgments as though such a relationship exists when in fact there is no relationship in the data. Thus, if individuals strongly believe that

costs within pools are correlated, these beliefs may overwhelm the data, and the individuals may overestimate within-pool zero predictors' coefficients. However, individuals might not overestimate the within-pool coefficients for several reasons. First, individuals will pay significant attention to the within-pool relationships, and therefore individuals will estimate the within-pool relationships accurately. Second, individuals have financial incentives to estimate coefficients accurately. And third, the data in my setting are relatively free from error, which will improve the accuracy of individuals' coefficient estimates.

The second way in which beliefs can affect judgment accuracy is by affecting the *variation* (that is, the random error) with which individuals apply their coefficient estimates. Estimates of relationships that are inconsistent with individuals' beliefs tend to have more random error than do estimates of relationships that are consistent with individuals' beliefs (Sniezek 1986; Luft and Shields 2001). Thus, because individuals without specific knowledge to the contrary will likely believe within-pool relationships are stronger or more likely to exist than across-pool relationships, within-pool coefficients will likely be estimated with less variation than across-pool coefficients. However, if individuals pay very little attention to across-pool coefficients, then the across-pool coefficients may be treated as though they are zero, and thus will exhibit little random error.

The above analysis leads to the following hypotheses in alternative form:

H1: Judgment accuracy will be higher (bias will be smaller and variation will not be larger) for within-pool non-zero predictors' coefficients than for across-pool non-zero predictors' coefficients.

H2: Judgment accuracy will be higher (bias will be smaller and variation will not be larger) for within-pool zero predictors' coefficients than for across-pool zero predictors' coefficients.

The existence of four types of coefficients (within- and across-pool; non-zero and zero correlations) means there are six possible pairwise comparisons. I have chosen to focus on the

two pairwise comparisons for several reasons. First, Hypothesis 3 presents an omnibus test of prediction accuracy, including all coefficients. Second, there is not a well-developed theoretical prediction for the four coefficient pairs that I do not test. Third, the pairs I test in hypotheses 1 and 2 allow for a clean test of whether differences in implicit coefficient estimation accuracy are due to differences in the overhead cost pool classification. That is, the untested pairs include a manipulation of both whether a coefficient is within-pool or across-pool *and* whether the environmental correlation is non-zero or zero. Thus, any test of those pairs is inherently a test of an interaction between overhead cost pool classification and cost behavior. Because of the lack of theoretical development and the fact that the emphasis in this portion of the paper is on overhead cost pool classification alone and not on its interaction with cost behavior, such a test would not be meaningful.

Hypothesis 3: Prediction accuracy of overhead costs

Hypotheses 1 and 2 make predictions about within-individual effects of overhead cost pool classification on implicit coefficient errors on *specific* predictors. Such evidence is important to understanding how overhead cost pool classification affects the accuracy of judgment in certain tasks (for example, estimating how much spoilage will change as a result of changing spending on one of the predictors). However, understanding how overhead cost pool classification affects performance on the overall cost prediction task requires understanding how cost behavior and *all* the individual coefficient estimates jointly affect cost prediction accuracy. Hypothesis 3 predicts how the interaction of overhead cost pool classification and between-individual manipulations of variable cost behavior affects a *set* of implicit coefficient errors and thus the overall accuracy of participants' predictions of overhead costs.

Location of stronger predictor

Individuals likely have beliefs that costs within pools behave similarly, and costs across pools behave dissimilarly, in the absence of strong cues to the contrary. However, empirical evidence shows that costs in different pools can be statistically significantly correlated (Anderson 1995; Ittner et al. 1997; Joshi et al. 2001). When the stronger predictor is across pools, individuals will devote less attention to the stronger across-pool relationship because it is inconsistent with their beliefs, and thus will underweight it. Further, individuals will have difficulty estimating the within-pool relationship because of its relatively low magnitude (Broniarczyk and Alba 1994). Because all of the implicit coefficients will be estimated less accurately, I expect that *ceteris paribus*, accuracy of individuals' predictions of overhead costs will be higher when the stronger predictor is within-pool than when it is across-pool.

Sign of relationships

Overhead costs can be either positively or negatively correlated. As noted above, typical recommendations for cost system design assume that costs within pools are positively correlated. However, empirical evidence demonstrates that overhead costs can exhibit either positive or negative correlations (e.g., Datar et al. 1993; Anderson 1995; Joshi et al. 2001). Prior research in psychology and accounting finds that estimation of negative relationships is inherently more cognitively difficult, and thus less accurate, than estimation of positive relationships (e.g., Brehmer 1974). Thus, even if individuals are aware of the *existence* of a negative coefficient relating predictor costs to the target cost, they may be less able to estimate accurately its *magnitude* than they would if the coefficient were positive. Therefore, I expect that *ceteris paribus*, the accuracy of individuals' predictions of overhead costs will be higher when relationships between overhead costs are positive than when they are negative.

Note that in this study, all the non-zero predictors are either positively or negatively correlated with the target; that is, there are no cases in which signs are mixed. While empirical evidence indicates that both positive and negative correlations occur simultaneously in practice (e.g., Datar et al. 1993), I chose to include only single-sign cases in this study for simplicity.

Interaction of relative strength and sign

I expect that the location of the stronger predictor and sign of the relationships among costs will ordinally interact to affect judgment accuracy. The effect of the location of the stronger predictor will be larger when relationships among costs are negative than when the relationships are positive. The expected interaction among the between-participant factors is presented graphically in Figure 5.

Insert Figure 5 here.

When relationships among costs are *positive*, the reduction in prediction accuracy (1 – 2 in Figure 5) that results from the stronger predictor being across-pool rather than within-pool will be relatively small. Individuals will allocate less attention to the stronger predictor when it is in a different pool from the target cost compared to when it is in the same pool. However, because the relationships among costs are positive, estimation is not as cognitively difficult, and so individuals will be able to estimate the implicit coefficient on the stronger predictor with relatively high accuracy. Further, the cognitive difficulty of estimating the implicit coefficient on the weaker, within-pool predictor will be relatively small because of the positive sign and the relatively high attention it receives because it is within-pool. Thus, the effect on prediction accuracy of the stronger predictor being across-pool rather than within-pool will be relatively small when relationships among costs are positive.

When relationships among costs are *negative*, the reduction in prediction accuracy (3 – 4 in Figure 5) that results from the stronger predictor being across-pool rather than within-pool will be relatively large. Individuals will allocate less attention to the stronger predictor when it is in a different overhead cost pool from the target cost compared to when it is in the same pool. The reduced attention individuals allocate to the stronger predictor will be insufficient to overcome the cognitive difficulty of estimating negative relationships, and individuals' estimates of the implicit coefficient on the stronger predictor will be less accurate. Further, the difficulty of estimating the implicit coefficient on the weaker, within-pool predictor will be exacerbated by the negative sign, despite the relatively high attention it receives because it is within-pool. Thus, neither implicit coefficient will be estimated accurately, and the effect on prediction accuracy of the stronger predictor being across-pool rather than within-pool will be larger when the relationships among costs are negative than when they are positive. That is, I predict the interaction will be of the form $3 - 4 > 1 - 2$. This prediction is consistent with evidence from Sniezek (1986), who found that the effect on judgment accuracy of inconsistency between individuals' beliefs and actual data is larger when relationships are negative than when they are positive.

H3: Overhead cost prediction accuracy is an interactive function of the location of the stronger predictor and the sign of relationship between the predictors and the target.

III. RESEARCH DESIGN

Research design

The hypotheses are tested using an experiment with a 2 x 2 x 2 x 2 x 2 mixed design. There are two within-participant factors. The first within-participant factor is whether a predictor is within the same pool as spoilage or across pools (that is, in a different pool from spoilage).

The second within-participant factor is whether a predictor has a non-zero coefficient or a zero coefficient relating the predictor to spoilage. There are three between-participant factors. The first between-participant factor is the location of the stronger predictor-target relationship, either within the same pool as spoilage or across pools. The second between-participant factor is the sign of the predictor-target relationships; they were all positive in one condition and all negative in the other. The third between-participants factor is the *label* of the costs. For half the participants, the predictors within the same pool as spoilage were labeled “inspection” and “testing,” while the predictors in the other pool were labeled “supervision” and “training.” For the other half of the participants, the labels were reversed, such that the predictors within the same pool as spoilage were labeled “supervision” and “training,” while the predictors in the other pool were labeled “inspection” and “testing.” This last factor was manipulated between participants to control for beliefs about relationships among costs with specific labels. There was not a main effect of label or a significant interaction between labels and other variables (all p values > 0.10); thus, this factor is omitted in all analyses reported below. The experimental design is shown in Table 1.¹²

Insert Table 1 here.

Participants

Eighty-one undergraduate business students enrolled in a cost-accounting class at a large public university volunteered to participate in the research. They were given course credit and \$5 for participating, and incentive-based pay from \$1 to \$10 based on their judgment accuracy.

¹² Participants’ responses were collected in five sessions. A control variable for session did not result in a significant main effect or interaction with other variables; thus, session is omitted in all analyses.

Three participants' responses were eliminated because their accuracy measures were influential outliers.¹³ The final sample thus included 78 participants.

Procedure and task

Participants were given a paper-and-pencil task in which they were asked to assume the role of a controller in a manufacturing firm. As part of their firm's operating budgeting process, they were to predict spoilage costs for 20 similar manufacturing plants for the next month.

Participants were asked to predict spoilage cost from the four other overhead costs (inspection, testing, supervisors, and training) in two overhead cost pools (quality control and supervision).

Either inspection or supervisors (depending on the label condition) was strongly predictive of spoilage, and the other was weakly predictive. The other two costs (testing and training) were not

predictive of spoilage. Each cost pool contained either the stronger or the weaker predictor and

one of the zero-correlation predictors. The coefficient relating the stronger (weaker) predictor to spoilage was 11.011 (10.461). The relationship between spoilage and the two predictive costs is:

$Spoilage = Constant \pm 1.01 \times Stronger \pm 0.46 \times Weaker + error$, where *Stronger (Weaker)* refers to the cost with the stronger (weaker) relationship, and *Constant* = -\$24,145 and \$104,125 in the positive and negative conditions, respectively, and *error* is zero-mean noise. The rows in Table 1 describe the specific model for each condition individually.

Participants were told spoilage costs were not the direct result of a single spending decision but resulted from, and were predictable from, other spending decisions made during the month. Participants were given a learning data set that provided the prior month's cost data by plant, reported in the two cost pools (see Appendix A, Exhibit 1).

¹³ Because mean squared error of predictions encompasses all potential causes of error, I use mean squared error to identify outliers. I regressed mean squared error on sign, location of the stronger predictor, and the interaction of sign and location to determine influential outliers using Cook's D. Two observations had values of Cook's D greater than $(4/n)$, a reasonable cutoff per Fox (1991), and were thus eliminated.

To reduce noise, I attempted to set participants' beliefs about the sign of relationships with the following wording for the positive and negative conditions, respectively:

“Higher spending on one (or more) of these activities could lead to higher levels of spoilage costs being recorded, because more care is taken to identify and discard defective units before they are sent to customers,” or “[h]igher spending on one (or more) of these activities could lead to lower spoilage costs by improving production quality.”

Participants also received a judgment data set, in which next month's budgeted levels of the predictor costs were given, with blanks to fill in predictions of expected spoilage costs (see Appendix A, Exhibit 2). Participants were allowed to use calculators.¹⁴ A series of follow-up questions included manipulation checks, GPA in accounting and statistics courses, questions about beliefs about the relationships among costs, and questions about participants' reasoning for examining particular relationships among costs.¹⁵

Experimental parameters

Within each sign condition, all data sets were statistically identical. I first generated standardized data that fit the data requirements for the research design. I then linearly transformed the standardized data to generate plausible levels of costs. The transformations were such that the means, variances and ranges of the two predictive costs (inspection and supervision) were identical to each other, thus eliminating any concerns about scale. The four columns of predictors were identical across experimental conditions; they differed only in their

¹⁴ One participant obviously used a statistical calculator, as he generated a very accurate model and reported its R^2 on his experimental materials. Results are not affected by the inclusion or exclusion of this participant's responses. In this task setting, the data set was small enough that complete testing of a multiple regression model was relatively easy, and the loss of efficiency caused by including all possible predictors was minimal. As noted above, the use of calculation aids will help individuals examine relations accurately and apply their decision models consistently. However, an important step in assuring accurate estimation is the choice of relationships to examine. The use of a calculation aid will not necessarily help in this part of the task.

¹⁵ A pair of follow-up questions asked for participants' estimates of the effect on spoilage of increased spending on two of the predictors. As noted in section V of this paper, I have yet to fully analyze the results from these questions, and so these questions are omitted from the analysis in this paper.

placement among cost pools.¹⁶ The only differences across any conditions were the values of spoilage, which differed between the negative and positive conditions by the use of a different linear transformation. Further, there were no statistically significant differences for the four predictor costs between the learning and judgment data sets. Thus, there is no concern about nonlinear relationships between spoilage and other costs due to mean differences between learning and judgment data.¹⁷ Parameters are described in Table 2.

Insert Table 2 here.

Measurement of variables

By regressing participants' predictions of the target on the levels of the predictors, I estimate their implicit coefficients on each predictor.¹⁸ I then estimate participants' implicit coefficient errors by subtracting the environmental coefficients from the implicit coefficients. Each participant's judgment model has a constant and coefficients for each of the four costs. I chose to treat an implicit coefficient as non-zero if it was significant at $p < 0.10$; otherwise I treated the implicit coefficient as zero.¹⁹ Thus, my models of participants' judgments include up to four coefficients. I also estimate the optimal or environmental model from the learning data set and generate the environmental coefficients, which represent the optimal coefficients participants

¹⁶ The predictive costs were always two columns away from spoilage, to control for any effects of proximity on judgment.

¹⁷ If means were different across learning and judgment data, participants might have assumed and applied nonlinear models, which would not be adequately captured by my linear modeling technique. This would result in inaccurate measures of their judgments.

¹⁸ I assume participants use a linear, additive model because the model used to generate the experimental data is linear and additive. Further, evidence from psychology research (Brehmer 1974; Dawes and Corrigan 1974) indicates that individuals typically use linear, additive models in subjective judgment.

¹⁹ My main results are unaffected by alternative choices of $p = 0.05$ or 0.10 .

could estimate. I compare participants' implicit coefficients to the environmental coefficients as described below.

To measure for an effect of overhead cost pool classification on *bias* (that is, differences between participants' within-pool and across-pool estimation errors) for Hypotheses 1 and 2, I compute a directional error difference measure equaling the signed error (negative signed error) on each coefficient when relationships are positive (negative). Thus, coefficients whose magnitudes are overestimated (underestimated) will have positive (negative) directional error, regardless of the sign of the relationships in the data. I then subtract the within-pool directional error from the across-pool directional error to create the directional error difference variables. I do this for the costs with non-zero coefficients for H1 and the costs with zero coefficients for H2 (*DirCE* and *DirZeroCE*, respectively). If either variable is negative (positive) and statistically significant, participants have on average underestimated (overestimated) across-pool coefficients more than within-pool coefficients. To test for an effect of overhead cost pool classification on *variation* for Hypotheses 1 and 2, I use the variance of the within-pool and across-pool coefficient errors.

For judgment accuracy in Hypothesis 3, I use several measures to test for differences in accuracy and explore reasons for those differences in accuracy. The main measure I use is the mean squared error of participants' predictions compared to optimal predictions made by the environmental model.²⁰ I also use three measures derived from the lens model to investigate how different sources of judgment error affect overall prediction accuracy.²¹ *Matching* is the

²⁰ Use of mean absolute error does not affect results.

²¹ Lens model studies typically include *achievement* as well, which is the correlation between participants' predictions and actual outcomes. Achievement is the product of three factors: matching, consistency and environmental predictability. In my study, environmental predictability is constant across conditions and is very

correlation between the predictions made by the environmental model and predictions made by each participant's model. Matching measures how accurately participants estimated the relative values of the coefficients for the four predictors. *Consistency* is the correlation between participants' predictions and the predictions made by their model. Consistency measures how consistently participants applied their implicit models. A final measure I use is *consensus*, or the averaged correlations of participants' predictions with predictions of other participants in the same experimental condition. Consensus is not a measure of individual accuracy but rather a measure of the similarity of participants' predictions to each other. However, this measure is important because low levels of consensus imply costly disagreement among individuals and reduced likelihood of reaching agreement about judgments and their effects on organization performance.

IV. RESULTS

Six participants from the final sample of 78 participants failed the post-task manipulation check question that asked them to identify which cost pool contained each of the five costs (the four predictors and spoilage). The failure of a post-task manipulation check may be due to failure of memory rather than one of attention or understanding. Further, the failure to recall the label on the overhead cost pool does not necessarily mean that the participant did not understand which individual cost accounts were classified together within each pool. Therefore, I retain the responses of the participants who failed the manipulation check; results are unchanged by the exclusion of these participants' data. Self-reported accounting and statistics GPA did not

close to its maximum value of 1.0; therefore, achievement is a redundant measure of judgment accuracy. For parsimony, I omit achievement from my study.

significantly affect any dependent variable or interact with any independent variable to affect any dependent variable ($p < 0.10$), and so are excluded from all analyses.

I also perform manipulation checks to ensure that participants attended to the cost behavior manipulations. I examine whether participants' implicit coefficients on the stronger predictor were larger than those on the weaker predictor, and whether participants' implicit coefficients carried the appropriate sign. Table 3 shows results of these manipulation checks. Panel A presents subjective coefficients for the non-zero predictors across the two sign conditions.²² Data are consistent with participants placing more weight on stronger predictors than on weaker predictors and implicitly identifying the signs of coefficients accurately.

Insert Table 3 here.

Panel B of Table 3 presents t-tests for mean differences between implicit coefficients on stronger and weaker predictors. Results in Panel B confirm that participants placed significantly more weight on stronger predictors than weaker predictors in both positive and negative conditions (p -values for nonzero < 0.001 for both signs). Thus, the results indicate that participants did not simply weight within-pool predictors and ignore across-pool predictors; rather, their pattern of weights indicates they paid attention to the relationships in the learning data set.

Hypothesis tests

Hypotheses 1 and 2

Hypotheses 1 and 2 predicted that participants will estimate within-pool implicit coefficients more accurately than across-pool implicit coefficients for non-zero and zero

²² Evidence on implicit coefficients on zero-correlation predictors is found in test of Hypotheses 1 and 2.

predictors, respectively. Panel A of Table 4 present tests for bias. The first row of Table 4 presents results for Hypothesis 1. Participants underweighted across-pool non-zero predictors more than within-pool non-zero predictors, consistent with beliefs that costs assigned within pools are more likely to be strongly related than are costs assigned into different pools. *DirCE* is significantly negative (mean = -0.18, $p = 0.005$). Thus, Hypothesis 1 is supported for bias. Results for Hypothesis 2 are shown in the second row of Panel A of Table 4. Participants did not implicitly overweight within-pool zero coefficients more than across-pool coefficients. *DirZeroCE* is not significantly different from zero (mean = 0.01, $p = 0.917$). Hypothesis 2 is not supported for bias.

Insert Table 4 here.

Panel B of Table 4 presents tests of Hypotheses 1 and 2 for variation. The variances of across-pool and within-pool non-zero predictors' coefficient errors are not statistically significantly different ($F = 1.01$; $p = 0.484$). Thus, Hypothesis 1 is supported for variation. For zero predictors' coefficients, the variance of the within-pool coefficient errors is significantly greater than that of the across-pool coefficient errors ($F = 1.61$; $p = 0.019$). Thus, Hypothesis 2 is not supported for variation.

Taken together, the results for Hypothesis 1 support the hypothesis for both bias and variation. Thus, I conclude that for non-zero predictors, judgment accuracy is higher for within-pool coefficients than for across-pool coefficients. That is, within-pool coefficients are estimated with less bias and no more variation than across-pool coefficients.²³ Taken together, the results for Hypothesis 2 do not support the hypothesis for either bias or variation. Thus, I conclude that

²³ Sensitivity tests indicate that there is no evidence that this effect varies across the cost behavior manipulations.

for zero-correlation predictors, judgment accuracy is not higher for within-pool coefficients than for across-pool coefficients. Specifically, bias is not statistically significantly different, and variation is higher for within-pool coefficients than for across-pool coefficients.

Hypothesis 3

Hypothesis 3 states an expectation that participants' prediction accuracy is a result of the interaction between the two manipulations of cost behavior (location of stronger relationship and sign), such that the stronger predictor being across pools will reduce overhead cost prediction accuracy more when relationships between the predictors and the target are negative than when they are positive. As noted above, I use four measures of accuracy. Table 6, Panel A presents descriptive statistics for the measures of accuracy. Cell-by-cell examination indicates accuracy is affected by the interaction of the cost behavior conditions as hypothesized. The decrease in accuracy between the across-pool and within-pool measures of accuracy is greater for the negative condition than for the positive condition.

Insert Table 6 here.

Panel B shows multivariate results of a MANOVA on the four measures of accuracy. As a group, they are significantly affected by the interaction of the location of stronger predictor and sign conditions ($p = 0.019$). Panel C shows the individual ANOVAs for each accuracy measure; all except consistency are significantly affected by the interaction (p values ≤ 0.075 ; p value for consistency = 0.181). Thus, H4 is partially supported. Note that Panel C of Table 6 indicates significant main effects of location of stronger predictor and sign on several measures of accuracy; however, due to the presence of the significant interaction effect, I do not interpret the main effects.

Interpretation of the results of Hypotheses 3 is as follows. For participants for whom the stronger predictor was across pools, those in the negative condition exhibited lower prediction accuracy than those in the positive condition. As revealed by their lower *matching* score, participants in this condition implicitly estimate the relative coefficients less accurately, indicating that their predictions result from less well-specified models. Thus, their predictions vary more from the optimal predictions, as indicated by higher *mean squared error*.²⁴ Finally, the participants in this cell use models that differ from each other more than those in other cells, as indicated by their lower levels of *consensus*.

Supplemental analysis

The above analysis demonstrates that overhead cost pool classification affects participants' estimation of implicit coefficients relating costs within and across pools, and the accuracy of their judgments of costs that result from their estimations. However, the results above do not specifically address *why* the effect occurs. Thus, I include analysis of experimental follow-up questions that provide some evidence on the question of why overhead cost pool classification affects coefficient estimation performance and judgment accuracy. Because these results rely on retrospective self-reports of cognitive processes, which have come under scrutiny and are commonly viewed as suspect (e.g., Nisbett and Wilson 1977), these results are considered supplementary.

²⁴ The interaction is significant for *consistency* ($p = 0.181$) but is directionally consistent with predictions, suggesting that participants in this cell applied their implicit coefficients to predictions with more variation than did participants in other conditions. Because participants were allowed to use calculators, they should have been able to apply their implicit coefficients equally consistently across cells. High levels of consistency are therefore to be expected across all experimental conditions if participants explicitly identified relationships and applied them using calculators. Therefore, the fact that consistency is almost significantly affected by the experimental conditions is surprising, and may indicate that participants in the negative, across-pool condition were less likely to explicitly identify coefficients, but rather used a variety of subjective strategies for applying implicit coefficients.

Table 7, Panel A presents the follow-up questions, which are intended to elicit responses from participants as to both *how* and *why* the overhead cost pool classification affected their estimation and prediction accuracy. Participants were asked to think about how they estimated the relationships, and then to allocate 100 points among six possible descriptions of their estimation strategies. Item C is intended to capture the likely beliefs that costs within pools are strongly correlated. Other items are intended to capture likely alternative beliefs and other foci of attention that may have resulted from the various conditions.

Insert Table 7 here.

Mean points allocated to each item are shown in Panel A.²⁵ Item C received more points than did any other item, indicating that on average, participants considered overhead cost pool classification an important factor in their cost prediction strategy because overhead cost pool classification likely describes overhead cost behavior. The difference between points allocated to Item C points allocated to all other items except Item A is statistically significant (Item A $p = 0.917$, all other p 's < 0.001). Note that item B, which offers responsibility accounting as an alternative reason for attending to overhead cost pool classification, received fewer points than did item C ($p = 0.001$, untabulated result).

It is not necessary that Item C receive significantly more points than any other item (and that the within-pool relationship receive strictly more attention than any other relationship) in order to support the hypothesis development in section 2. It is sufficient that some participants allocate a non-trivial amount of attention to the relationship and that that attention affect the

²⁵ Untabulated results indicate that points allocated to the items do not vary among the cost behavior conditions. Thus, the point allocations likely do not reflect effects of data on prediction strategies.

accuracy of those participants' judgments. By construction, the objectively optimal strategy is to ignore the overhead cost pool classification and to use all the data independently of its pool location. Thus, any deviation from that strategy in the direction suggested by the overhead cost pool classification that reduces participants' judgment accuracy supports the hypothesis development.

If the overhead cost pool classification affected the allocation of participants' attention to relationships in the data, and the allocation of attention affected the accuracy of predictions, then there should be an effect of the participants' prediction strategies (as measured by the self-reported points allocation) on the error of participants' predictions.²⁶ This effect should be conditional on the location of the stronger predictor. When the stronger predictor is within the same pool as the target, then the prediction task is relatively easy. Given the magnitude of the stronger relationship and the lack of noise in the data, even moderate attention paid to the within-pool predictor should enable participants to estimate the coefficient on the strong predictor fairly accurately. Thus, strategies that involve either looking first at the within-pool relationship (items B and C) or looking at all relationships regardless of pool location (items A and E) should be associated with lower error. A strategy of first looking at the across-pool relationship (item D), however, would lead to higher error, as participants would allocate the majority of their attention toward the weaker predictor. Thus, points allocated to items A, B, C and E should be negatively correlated with prediction error, while points allocated to item D should be positively correlated with prediction error.

²⁶ Untabulated results indicate that correlations among points allocated to the post-experimental questions and *DirCE* are directionally consistent with the discussion below relating strategies to overall prediction error (MSE and MAE), but are not statistically significant.

Conversely, when the stronger predictor is in the opposite pool as the target, a strategy of allocation attention first to predictors within the same pool as the target and later to predictors in the other pool will result in higher error. Thus, for across-pool participants, prediction error should be positively correlated with points allocated to those items (B and C) that indicate prediction strategies that allocate attention first to predictors in the same pool as the target. A strategy of allocating attention to across-pool relationships (item D) or to all relationships regardless of pool location (items A and E) will allow participants to identify the stronger predictor in the opposite pool without first having allocated much attention to the within-pool relationship, and so points allocated to these items should be negative correlated with error.

Table 7, Panel C presents correlations among the points allocations and prediction error. For the within-pool participants, prediction error measures do not significantly correlate with points allocations. However, several of the correlations are quantitatively in the predicted direction. Correlations between both error measures and Item C are negative.²⁷ Correlations between error measures and item D are positive and nearly significant for MAE ($p = 0.11$, two-tailed). Thus, those subjects whose prediction strategies included early allocation of attention to predictors with the same pool as the target (which includes the strong predictor) made predictions with less error, and those who allocated initial attention to the predictors in the other pool (including the weak predictor) made predictions with more error.

For across-pool participants, the results are somewhat stronger. Points allocated to item C are positively correlated with both error measures (p values for MSE and MAE both < 0.10) and points allocated to Item A are negatively correlated with both error measures (p value for MSE $<$

²⁷ It is possible that the range of points allocated to items affects the correlations. The ranges for items are: A: 100; B: 60; C: 100; D: 75; E: 80; and F: 40. Thus, items A and C, which have the greatest ranges, also have statistically significant correlations with error measures for the across-pool participants, while other items have no significant correlations.

0.10; p value for MAE = 0.12). Thus, those subjects whose prediction strategies included early allocation of attention to predictors with the same pool as the target (which includes the weak predictor) made predictions with more error, and those who allocated initial attention to the predictors in the other pool (including the strong predictor) made predictions with less error.

Though the results are not consistently statistically significant, the supplemental analysis overall is consistent with the theoretical development in section II above. The evidence is consistent with prediction error resulting in part from participants' allocation of attention among various predictors, and with participants allocating attention in part based on the classification of costs within pools.

V. DISCUSSION AND CONCLUSION

Discussion

This study provides evidence that overhead cost pool classification affects the accuracy of judgments about cost behavior by users of accounting systems. I demonstrate that while individuals do pay significant attention to the data they are given, they estimate coefficients within overhead cost pools more accurately than they do coefficients across overhead cost pools, and there is evidence consistent with the effect being robust to differences in the location of a stronger vs. a weaker predictive relationship and the sign of the coefficients. I also show that individuals underestimate the magnitude of across-pool coefficients more than they do for within-pool coefficients. This pattern is consistent with empirical evidence found in Joshi et al. (2001), in which the magnitude of relationships among costs assigned into different cost pools was underestimated by managers. While there are likely other causes for the estimation errors exhibited by managers at those firms, I have demonstrated that overhead cost pool classification is a plausible cause of the errors.

I have also shown that the accuracy of participants' cost predictions depends on an interaction between the pool location of the stronger predictive relationship and the sign of coefficients. Participants facing negative relationships between costs and stronger predictors located across cost pools exhibited greater error and lower consensus than did other participants. More broadly, my results provide evidence on tradeoffs organizations face when designing cost accounting systems. Organizations' overhead cost pool classification decisions should be informed by judgment-related effects as well as the potential for product costing errors as described by prior literature.

This study complements the research on the choice of cost drivers and the degree of aggregation of costs into pools by addressing judgment effects of cost system design. The finding that overhead cost pool classification itself communicates information to individuals also contributes to literature on classification in financial statements (e.g., Hopkins 1996) and suggests future directions for research into other areas of accounting classification: for example, variance analysis and other responsibility accounting methods, the placement of subtotals on financial statements or earnings announcements, and the format of pro-forma earnings reconciliations may all be areas of fruitful study. Further, the boundary conditions or variables that modify the effects of classification decisions are not yet known. Whether and how the effects of classification decisions vary by characteristics of the task, individual, and other details of the behavior of the classified data have yet to be studied.

This study also has implications for both practitioner and classroom discussions of cost system design. Instruction and discussions of cost system design would likely benefit from an explication of the implicit assumption of similar (dissimilar) cost behavior within (across) pools, and the potential effects on individuals' judgment. This may be especially important in two

particular cases. First, refined costing systems such as ABC and the German cost accounting system known as *Grenzplankostenrechnung* (see, e.g., Sharman 2003) contain many cost pools or cost centers. My study implies that increasing the number of cost pools within a cost system also increases the possibility that cost relationships across different pools will receive less attention and be misestimated. This can affect the likelihood that individuals will be able to use their costing system to identify possible improvements or areas of concern. Second, the introduction of sophisticated enterprise resource planning and information systems has increased some managers' trust in, and reliance on the output of, their accounting systems. For example, one manager stated that after his firm's implementation of a new enterprise planning system, individuals within the firm trust their system's output because "the data is the data" (Anonymous 2006). Thus, managers may have less inclination to examine what is "inside the black box" of classifications reported by a sophisticated information system. In these cases, the effect of cost system design on judgments of cost behavior may be especially pernicious.

Limitations

This study is subject to several limitations. First, the usual concerns about using student participants in experimental settings apply. Thus, to the extent that the results are not caused by cognitive processes common to most individuals educated in basic cost accounting, or may be affected by experience or knowledge, my results may not be representative of those that might be obtained using professionals performing familiar tasks. However, the fact that important across-pool relationships were grossly underestimated by experienced managers as noted in Joshi et al. (2001) provides corroborating evidence that the effects noted here exist outside the laboratory. Second, my study used a reporting format designed primarily for internal validity rather than for efficient managerial use. There is likely great variation in how individual overhead costs and overhead cost pools are reported to individuals, and details of the reports that differ from the

choices made here may be important to practitioners. Also, I examine a setting with few potential relationships and only four of many possible combinations of sign and relative strength of coefficients. There may be important coefficient structures (for example, mixed signs or larger numbers of potential relationships) that are yet to be examined.

Appendix A: Experimental instrument

Background

Assume you are the controller of a manufacturing firm. One of the important parts of your job is to *predict future costs* as part of your firm's short-term budgeting process. One cost that you must predict for the plants in your firm is the *monthly reported spoilage cost* – that is, the cost of products that are discarded because they do not pass inspection. This cost must be predicted because it is not an amount that plant managers directly decide to spend, but is the result of things that happen during the month.

Some (not necessarily all) of the other activities at the plant influence the amount of spoilage each month. The main activities in the plant other than direct production are **inspection, testing, supervision, and training**. *Higher spending* on one (or more) of these activities could lead to *higher levels of spoilage costs* being recorded, because more care is taken to *identify and discard defective units* before they are sent to customers. No formal statistical analysis of the relationships among these costs has been performed, so it is not yet known which, if any, of the other activity costs are good predictors of spoilage costs.

The plants in your region are all comparable in terms of workforce, equipment and product mix, so the relationships, if any, between inspection, supervision, training and spoilage are very similar across plants. Also, wages and input prices are similar across plants, so differences in spending levels represent differences in activity levels, not differences in prices or wages.

In order to help you predict next month's spoilage costs, your plants' cost accountants have provided the information on **Exhibit 1** from the plants' cost accounting systems. The information includes the last month's spoilage costs. Also included in Exhibit 1 are the other costs that could be helpful in predicting spoilage costs. Last month's costs are typical for your firm, and should offer a good basis for predicting next month's costs.

Note that the firm's cost accounting system reports all these costs in two *cost pools* for each plant:

1. The *Quality Control pool* includes *inspection* and *testing* costs. This pool captures the costs of ensuring that no defective products go out to customers.
2. The *Supervision pool* includes wages and benefits for *supervisors* and the costs of *training programs*, and also *spoilage* costs. This pool captures the costs of preventing poor quality work by ensuring the line employees are well trained and have adequate supervision.

Costs within each pool are totaled, and only the pool total is regularly reported to most managers. The pool total is used to allocate costs to individual products and customers.

Please examine Exhibit 1 and *learn as much as you can* about predicting spoilage costs from the other costs provided. When you are finished, please continue on to the next page.

Cost prediction

When you are finished examining Exhibit 1, please turn to **Exhibit 2**. Exhibit 2 contains partial budgets for the region's plants, as prepared by the plants' managers. The amounts included on Exhibit 2 include the monthly quality control and supervision costs the plant managers have budgeted for next month.

Based on your conversations with the cost accountants and plant managers and your review of your plants' production plans, you believe the budgeted amounts on Exhibit 2 are reasonable. You also believe that the relationships between these costs and spoilage costs are likely to be the same in the next month as they were last month.

Please use the budgeted costs on Exhibit 2 to **predict monthly spoilage costs** for next month based on what you have learned from Exhibit 1. Feel free to use a calculator.

It is important that you try to predict costs as accurately as possible. If your predictions are inaccurate, the firm's future profitability and your performance evaluations will suffer.

Recall also that your pay is based on your predictions. Your predictions will be compared to the best predictions that can be made from the data provided. The closer your predictions to the best possible predictions, the higher your pay will be.

The plant names on Exhibit 2 are disguised. It is not a good strategy to try to "match" one plant on Exhibit 2 with one or two plants on Exhibit 1. It is a ***better strategy*** to try to ***learn relationships among the costs*** and use these relationships to predict the spoilage costs.

Please complete the 20 spoilage cost predictions now.

Once you are finished with your predictions, please turn to the next page.

[Note: this is the instrument for the following conditions: positive correlations, stronger predictor is within-pool, spoilage is in the supervisory costs pool.]

Advice to top management

Now that you have learned about the relationships among the various costs and have made your predictions, top management has asked you for your advice on a decision. Because of recent concerns about product quality, top management would like to ***increase its spending on quality-improvement initiatives***. However, top managers are unsure about how to make the greatest impact on quality-related costs. Thus, they are considering two different options. Note that these two options are ***mutually exclusive*** – that is, ***management will pursue only one of them, but not both***.

For help in deciding which option to pursue, management has asked you for an estimate of the effect of each of the two options. ***Your estimate should be based on the relationships you have learned from Exhibit 1***. Please feel free to refer to Exhibit 1 when making these estimates.

Option 1: What would be the change in spoilage costs if management gave ***each plant an additional \$10,000 per month to spend on inspection***? There would be no change in any other spending levels.

Each plant's monthly spoilage costs would be expected to (circle one) **increase** or **decrease** by about \$_____.

Option 2: What would be the change in spoilage costs if management gave ***each plant an additional \$10,000 per month to spend on supervisors***? There would be no change in any other spending levels.

Each plant's monthly spoilage costs would be expected to (circle one) **increase** or **decrease** by about \$_____.

Thank you very much. Now, please ***return this packet*** to the researcher, and ***pick up your second packet***.

Exhibit 1: Plants' cost data from cost accounting system

Plant	Supervisory Costs cost pool				Quality Control cost pool		
	Details			Pool	Details		Pool
	Supervisors	Training	Spoilage	Total	Inspection	Testing	Total
1	\$ 16,800	\$ 14,000	\$ 12,700	\$ 43,500	\$ 39,000	\$ 32,000	\$ 71,000
2	28,000	18,300	25,100	71,400	46,500	23,000	69,500
3	62,900	18,400	61,200	142,500	29,000	27,200	56,200
4	36,400	16,200	21,700	74,300	28,000	31,300	59,300
5	35,000	18,000	21,500	74,500	32,000	27,500	59,500
6	58,500	12,500	60,700	131,700	59,500	26,000	85,500
7	51,700	10,200	39,500	101,400	32,500	27,500	60,000
8	48,500	19,000	26,800	94,300	16,900	23,100	40,000
9	29,000	17,000	41,400	87,400	55,100	24,400	79,500
10	49,000	19,600	47,700	116,300	36,000	28,100	64,100
11	30,000	15,000	15,400	60,400	37,000	21,500	58,500
12	21,600	15,900	13,600	51,100	37,300	30,600	67,900
13	42,000	13,800	44,300	100,100	36,800	23,700	60,500
14	56,800	18,600	53,400	128,800	50,000	24,500	74,500
15	38,500	14,200	29,500	82,200	25,300	32,500	57,800
16	64,300	13,000	69,300	146,600	53,600	29,100	82,700
17	43,700	15,300	26,100	85,100	48,500	29,000	77,500
18	22,500	10,200	15,300	48,000	37,500	25,600	63,100
19	20,100	14,100	12,900	47,100	23,600	25,700	49,300
20	16,600	16,200	12,700	45,500	45,500	24,900	70,400

Exhibit 2: Plants' partial budgeted cost data

<u>Plant</u>	<u>Supervisory Costs cost pool</u>			<u>Quality Control cost pool</u>	
	<u>Details</u>			<u>Details</u>	
	<u>Supervisors</u>	<u>Training</u>	<u>Spoilage</u>	<u>Inspection</u>	<u>Testing</u>
A	\$ 57,500	\$ 14,000	_____	\$ 55,500	\$ 26,000
B	35,000	16,300	_____	23,500	32,300
C	20,500	14,200	_____	34,200	22,300
D	30,000	13,000	_____	35,000	27,900
E	20,500	14,800	_____	47,500	24,800
F	40,000	15,000	_____	39,900	30,100
G	17,500	13,200	_____	22,000	30,500
H	53,000	14,500	_____	21,500	34,200
I	22,000	17,500	_____	37,400	27,100
J	30,900	11,100	_____	33,500	27,500
K	40,000	17,900	_____	20,700	23,800
L	26,900	16,800	_____	19,000	30,200
M	46,100	13,900	_____	59,400	31,400
N	45,200	10,200	_____	18,900	29,000
O	36,000	11,000	_____	64,100	29,000
P	29,600	16,800	_____	49,000	31,300
Q	24,200	15,000	_____	42,500	33,700
R	41,800	13,100	_____	33,000	20,000
S	64,900	16,100	_____	58,000	21,700
T	56,700	14,800	_____	23,000	31,100

References

- Anderson, S. W. (1995). Measuring the impact of product mix heterogeneity on manufacturing overhead cost. *The Accounting Review*, 70(3), 363.
- Anonymous (2006). The data is the data. *Business Performance Management Magazine*. Vol. 4, Iss. 1, p. 36
- Banker, R. D. and S. M. Datar. (1989). Sensitivity, precision, and linear aggregation of signals for performance evaluation. *Journal of Accounting Research* (Spring): 21-39.
- Banker, R., and H. Johnston (1993). An empirical study of cost drivers in the U.S. airline industry. *The Accounting Review*, 68(3), 576.
- Banker, Rajiv D., G. Potter, and R. Schroeder. (1995). An empirical analysis of manufacturing overhead cost drivers. *Journal of Accounting & Economics* 19 (February):115-137.
- Billman, D., B. Bornstein, and J. Richards. (1992). Effects of expectancy on assessing covariation in data: "Prior belief" versus "meaning." *Organizational Behavior and Human Decision Processes* 53, 74-88.
- Brehmer, B. (1974). Hypotheses about relations between scaled variables in the learning of probabilistic inference tasks. *Organizational Behavior and Human Decision Processes*, 11(1), 1-27
- Briers, M., C. Chow, N. Hwang and P. Lockett. (1999). The effects of alternative types of feedback on product-related decision performance: a research note. *Journal of Management Accounting Research* 11, 75-92
- Bromwich, M., and C. Hong. (1999). Activity-based costing systems and incremental costs. *Management Accounting Research* 10 (1), pp. 39-60
- Broniarczyk, S. M. and J. Alba. (1994), Theory versus data in prediction and correlation tasks. *Organizational Behavior and Human Decision Processes*, 57, 117-139.
- Christensen, J. and J. Demski. (1995). The classical foundations of "modern" costing. *Management Accounting Research*, 6(1), 13.
- Christensen, J. and J. Demski. (1997). Product costing in the presence of endogenous subcost functions. *Review of Accounting Studies* 2 (1), 65-87.
- Cooper, R. and R. Kaplan (1999) *Design of Cost Management Systems: 2nd Edition* New Jersey: Prentice Hall

- Cooper, R., R. Kaplan, L. Maisel, E. Morrissey, and R. Oehm. (1992). *Implementing Activity-Based Costing: Moving from Analysis to Action*. Montvale, NJ: Institute of Management Accountants
- Datar, S., M. Gupta, (1994). Aggregation, specification and measurement errors in product costing. *The Accounting Review*, 69(4), 567.
- Datar, S. M., S. Kekre, T. Mukhopadhyay, and K. Srinivasan. (1993). Simultaneous estimation of cost drivers. *The Accounting Review*, 68(3), 602.
- Dawes, R. M. and B. Corrigan (1974). Linear models in decision making. *Psychological Bulletin*, 81, pp. 95-106
- Drake, A., S. Haka, and S. Ravenscroft (1999). Cost system and incentive structure effects on innovation, efficiency and profitability in teams. *The Accounting Review* 74 (3): 323-345.
- Foster, G. and M. Gupta. (1990). Manufacturing overhead cost driver analysis. *Journal Of Accounting & Economics*, 12(1-3), 309.
- Fox, J. (1991). *Regression Diagnostics*. Thousand Oaks, CA: Sage Publications.
- Garrison, R., E. Noreen, and P. Brewer. (2006). *Managerial Accounting*. 11th ed. New York: McGraw-Hill.
- Gupta, M., and R. King (1997). An experimental investigation of the effect of cost information and feedback on product cost decisions. *Contemporary Accounting Research*, 14 (1) 99-127.
- Hopkins P. (1996). "The Effect of Financial Statement Classification of Hybrid Financial Instruments On Financial Analysts' Stock Price Judgments" *Journal of Accounting Research* (Supplement): 33-50
- Horngrén, C., S. Datar, G. Foster. (2002) *Cost Accounting: A Managerial Emphasis*, 11th ed. New Jersey: Prentice Hall
- Ittner, C., Larcker, D, and M. Meyer. (2003). Subjectivity and the weighting of performance measures: Evidence from a balanced scorecard. *The Accounting Review*, 78 (3), 725
- Ittner, C., D. Larcker, and T. Randall. (1997). The activity-based cost hierarchy, production policies and firm profitability. *Journal of Management Accounting Research* 9: 143-162.
- Joshi, S., R. Krishnan, and L. Lave. (2001). Estimating the hidden costs of environmental regulation. *Accounting Review* 76 (2), 171-198
- Kaplan, R. S., and R. Cooper. (1998). *Cost and Effect: Using Integrated Cost Systems to Drive Profitability and Performance*. Boston: Harvard Business School Press.

- Kaplan, R. S. and D. P. Norton. (1996). *The Balanced Scorecard: Translating Strategy into Action*. Boston: Harvard Business School Press.
- Kaplan, R. S. and D. P. Norton. (2001). *The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment*. Boston, MA: Harvard Business School Press.
- Luft, J. and M. Shields. (2001). Why does fixation persist? Experimental evidence on the judgment performance effects of expensing intangibles. *The Accounting Review* 76 (3): 561-587.
- Markman, A.B., & Gentner, D. (2001). Thinking. *Annual Review of Psychology*, 52, 223-247.
- McKinnon, S. and Bruns, W. (1992). *The Information Mosaic*. Boston: Harvard Business School Press
- Muchinsky, P. and A. Dudycha. (1974). The influence of a suppressor variable and labeled stimuli on multiple cue probability learning. *Organizational Behavior and Human Decision Processes*, 12(3), 429-444.
- Muchinsky, P. and A. Dudycha. (1975). Human inference behavior in abstract and meaningful environments. *Organizational Behavior and Human Decision Processes*, 13(3), 377-391.
- Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, 92, 289-316.
- Nisbett, R., and T. Wilson. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review* 84, 231-259.
- Noreen, E. (1991). Conditions under which activity-based cost systems provide relevant costs. *Journal of Management Accounting Research* 3, p. 159-168.
- Roth, H. and Borthick, A. F. (1991). Are you distorting costs by violating ABC assumptions? *Management Accounting* 73 (5) 39-42.
- Sharman, P. A. (2003). Bring on German cost accounting. *Strategic Finance* 85 (6) 30-38.
- Snizek, J.A. (1986). The role of labels in cue probability learning tasks. *Organizational Behavior and Human Decision Processes* 38, 141-161.
- Yamauchi, T., & Markman, A.B. (2000). Inference using categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 776-795.

Table 1: Experimental design

Quality control pool
Location of stronger relationship and sign

Within-pool

(1) *Positive*

(2) *Negative*

Across-pool

(3) *Positive*

(4) *Negative*

		<u>Quality control pool</u>		
		<u>Spoilage</u>	<u>Inspection</u>	<u>Testing</u>
(1)	<i>Positive</i>	--	1.01	0.0
(2)	<i>Negative</i>	--	-1.01	0.0
(3)	<i>Positive</i>	--	0.46	0.0
(4)	<i>Negative</i>	--	-0.46	0.0

		<u>Supervision pool</u>	
		<u>Supervisors</u>	<u>Training</u>
		0.46	0.0
		-0.46	0.0
		1.01	0.0
		-1.01	0.0

Supervision pool

Within-pool

(5) *Positive*

(6) *Negative*

Across-pool

(7) *Positive*

(8) *Negative*

		<u>Supervision pool</u>		
		<u>Spoilage</u>	<u>Supervisors</u>	<u>Training</u>
(5)	<i>Positive</i>	--	1.01	0.0
(6)	<i>Negative</i>	--	-1.01	0.0
(7)	<i>Positive</i>	--	0.46	0.0
(8)	<i>Negative</i>	--	-0.46	0.0

		<u>Quality control pool</u>	
		<u>Inspection</u>	<u>Testing</u>
		0.46	0.0
		-0.46	0.0
		1.01	0.0
		-1.01	0.0

The entries in the cells are the coefficients relating each predictor cost to spoilage. Each row, representing a between-participant cost behavior condition, can be read as the specific environmental model for that cost behavior condition, with the addition of a constant. The constant varied between positive and negative sign conditions (-\$24,145 and \$104,125, respectively). For example, for participants receiving the cost behavior described in row (1), Spoilage = -\$24,145 + 1.01 x Inspection cost + 0.46 x Supervisors cost. For participants receiving the cost behavior described in row (6), Spoilage = \$104,125 - 1.01 x Supervisors cost - 0.46 x Inspection cost.

The first between-participant factor was the location of the stronger predictor-target relationship (correlation = |0.91; coefficient = |1.01|), either within the same pool as spoilage or across pools; the other relationship was the weaker one (correlation = |0.46|; coefficient = |0.46|). The second between-participant factor was the sign of the predictor-target relationships; they were all either positive or negative. The third between-participants factor was the label of the costs. For half the participants, the predictors within the same pool as spoilage were labeled “inspection” and “testing,” while the predictors in the other pool were labeled “supervision” and “training.” For the other half of the participants, the labels were reversed, such that the predictors within the same pool as spoilage were labeled “supervision” and “training,” while the predictors in the other pool were labeled “inspection” and “testing.” The first within-participant factor is whether a predictor is within the same pool as spoilage or across pools. The second within-participant factor is whether a predictor has a non-zero coefficient or a zero coefficient.

Table 2: Experimental parameters

Predictor	Dataset	Mean	P value of t-test for mean difference	Std. Deviation	P value of Levene's test for s.d. difference
Stronger	Learning	38,595	0.719	15,457	0.482
	Judgment	36,915		13,756	
Weaker	Learning	38,480	0.702	11,345	0.177
	Judgment	36,880		14,713	
Zero Coefficient	Learning	26,860	0.251	3,170	0.307
Within-pool	Judgment	28,195		4,022	
Zero Coefficient	Learning	15,475	0.202	2,763	0.189
Across-pool	Judgment	14,460		2,134	

Environmental model: Spoilage = Constant \pm 1.01*Stronger \pm 0.46*Weaker
where Constant = -\$24,125 in the positive condition and \$104,125 in the negative condition

Note: I use Levene's test because the underlying data are known to be non-normal, and the F-test is sensitive to departures from normality.

This table demonstrates that the data used for the learning data sets are not statistically significantly different from those in the judgment data set.

Table 3: Manipulation checks***Panel A: Participants' coefficients***

Sign		Stronger predictor	Weaker predictor
Positive	Mean	0.65	0.14
	N	41	41
	Std. Deviation	0.41	0.33
Negative	Mean	-0.65	-0.06
	N	37	37
	Std. Deviation	0.81	0.36

This panel shows the mean implicit coefficients on non-zero predictors in participants' subjective prediction models by sign condition.

Panel B: T-tests

Sign	Comparison	Mean	Std. Error Mean	t	df	Sig. (2-tailed)
Positive	Stronger predictor - Weaker predictor	0.52	0.10	5.378	40	0.000
Negative	Stronger predictor - Weaker predictor	-0.59	0.09	-6.00	36	0.000

This panel shows t-tests for differences between participants' implicit coefficients to test whether participants attended to data.

Table 4: Hypotheses 1 and 2**Panel A: Bias**

H	Dependent variable	n	(A) Mean across-pool implicit coefficient estimation error	(B) Mean within-pool implicit coefficient estimation error	Dependent variable (A) – (B) = (C) Mean difference (across-pool minus within- pool)	Paired samples t- value	Sig. (2- tailed)
1	<i>DirCE</i>	78	-0.45	-0.27	-0.18	2.867	0.005
2	<i>DirZeroCE</i>	78	0.12	0.11	0.01	-0.104	0.917

This panel presents descriptive detail and differences for the measures used to test Hypotheses 1 and 2. Columns A and B present mean across-pool and within-pool implicit coefficient errors, respectively. Column C presents the difference between across-pool and within-pool implicit coefficient errors. The differences presented in column C are the dependent variables of interest used in testing Hypotheses 1 and 2.

Directional error is the signed error (opposite of signed error) for participants in the positive (negative) condition. Positive (negative) directional error indicates overweighting (underweighting). Non-zero (zero) coefficient indicates that the errors represented in the row are implicit coefficient errors on predictors with non-zero (zero) environmental coefficients relating the predictor to the target cost.

Table 4: Hypotheses 1 and 2, continued**Panel B: Variation**

Hypothesis	Coefficient type	n	(A)	(B)	Variance ratio	Sig. (2-tailed)
			Across-pool implicit coefficient estimation error variance	Within-pool implicit coefficient estimation error variance		
1	Non-zero	78	0.172	0.171	1.01	0.484
2	Zero	78	0.362	0.584	1.61	0.019

This panel presents descriptive detail and differences for the measures used to test Hypotheses 1 and 2. Columns A and B present mean across-pool and within-pool implicit coefficient errors, respectively. Column C presents the difference between across-pool and within-pool implicit coefficient errors. The differences presented in column C are the dependent variables of interest used in testing Hypotheses 1 and 2.

Directional error is the signed error (opposite of signed error) for participants in the positive (negative) condition. Positive (negative) directional error indicates overweighting (underweighting). Non-zero (zero) coefficient indicates that the errors represented in the row are implicit coefficient errors on predictors with non-zero (zero) environmental coefficients relating the predictor to the target cost.

Table 5: Sensitivity analysis: across-pool minus within-pool implicit coefficient errors by cost behavior condition: Each cost behavior condition contrasted with the mean of the other three cost behavior conditions

Between-participant condition (location of stronger predictor, sign) compared to the mean of the other three conditions	Contrast estimate	Standard error	Sig.	Partial eta squared
Within-pool, positive	-0.24	0.45	0.596	0.004
Across-pool, positive	0.13	0.45	0.779	0.001
Within-pool, negative	0.16	0.46	0.728	0.002
Across-pool, negative	-0.05	0.47	0.913	0.000

This table shows the result of specified contrast tests for a mixed ANOVA. The within-participants effect is the directional implicit coefficient estimation error for the across-pool predictor minus the directional implicit coefficient estimation error for the within-pool predictor. The between-participants effects are the two manipulations of location of stronger predictor and sign of relationships.

Directional error is the signed error (opposite of signed error) for participants in the positive (negative) condition. Positive (negative) directional error indicates implicit overweighting (underweighting) of a predictor.

The contrast estimate is the result for a contrast that specifies the difference between the mean of the within-participant effect for the cost behavior condition and the mean of the within-participant effects for all other data conditions.

Table 6: Hypothesis 3
Panel A: Descriptive statistics

	Sign	Stronger Predictor	Mean	Std. Deviation	N
Mean Squared Error	Positive	Within pool	223,234,087	212,028,568	21
		Across pools	169,395,071	136,904,061	20
		Total	196,971,152	179,229,418	41
	Negative	Within pool	226,601,499	277,416,666	19
		Across pools	367,097,536	296,452,145	18
		Total	294,950,922	291,631,408	37
	Total	Within pool	224,833,607	242,027,520	40
		Across pools	263,043,607	244,972,613	38
		Total	243,448,735	242,642,256	78
Matching	Positive	Within pool	.75	.36	21
		Across pools	.84	.18	20
		Total	.79	.29	41
	Negative	Within pool	.76	.40	19
		Across pools	.36	.67	18
		Total	.56	.58	37
	Total	Within pool	.75	.38	40
		Across pools	.61	.53	38
		Total	.68	.46	78
Consistency	Positive	Within pool	.85	.22	21
		Across pools	.85	.14	20
		Total	.85	.18	41
	Negative	Within pool	.82	.25	19
		Across pools	.65	.42	18
		Total	.74	.35	37
	Total	Within pool	.84	.23	40
		Across pools	.75	.32	38
		Total	.80	.28	78
Consensus	Positive	Within pool	.48	.32	21
		Across pools	.52	.16	20
		Total	.50	.25	41
	Negative	Within pool	.49	.20	19
		Across pools	.12	.28	18
		Total	.31	.30	37
	Total	Within pool	.48	.26	40
		Across pools	.33	.30	38
		Total	.41	.29	78

Sign is the sign of relationships among costs.

StrongPredictor refers to the location of the stronger relationship.

Mean squared error is the mean squared error of participants' predictions compared to optimal predictions.

Matching is the correlation between the predictions made by the environmental model and predictions made by the participants' model.

Consistency is the correlation between participants' predictions and the predictions made by their model.

Consensus is the averaged correlations of participants' predictions with predictions of other participants in the same experimental condition.

Table 6: Hypothesis 3, continued

Panel B: Multivariate ANOVA

Effect	Wilks' Lambda	F(a)	Hypothesis df	Error df	Sig.
Intercept	0.03	552.23	4	71	0.000
Sign	0.84	3.281	4	71	0.016
StrongPredictor	0.86	2.87	4	71	0.029
Sign * StrongPredictor	0.85	3.17	4	71	0.019

a Exact statistic

This panel presents the results of a MANOVA, with the following variables:

Independent variables:

Sign is the sign of relationships among costs.

StrongPredictor refers to the location of the stronger relationship.

Dependent variables:

Mean squared error is the MSE of participants' predictions compared to predictions derived from the environmental model.

Matching is the correlation between the predictions made by the environmental model and predictions made by the participants' model.

Consistency is the correlation between participants' predictions and the predictions made by their model.

Consensus is the averaged correlations of participants' predictions with predictions of other participants in the same experimental condition.

Table 6: Hypothesis 3, continued**Panel C: Individual ANOVAs**

Source	Dependent Variable	Mean Square	F	Sig.
Corrected Model	Mean Squared Error	1.33E+17	2.380	.076
	Matching	.856	4.581	.005
	Consistency	.177	2.415	.073
	Consensus	.637	10.477	.000
Intercept	Mean Squared Error	4.73E+17	84.604	.000
	Matching	35.388	189.326	.000
	Consistency	48.811	664.833	.000
	Consensus	12.603	207.276	.000
Sign	Mean Squared Error	1.96E+17	3.516	.065
	Matching	1.072	5.736	.019
	Consistency	.268	3.652	.060
	Consensus	.737	12.115	.001
StrongPredictor	Mean Squared Error	3.65E+16	.653	.422
	Matching	.458	2.452	.122
	Consistency	.154	2.096	.152
	Consensus	.520	8.553	.005
Sign * StrongPredictor	Mean Squared Error	1.84E+17	3.284	.074
	Matching	1.167	6.243	.015
	Consistency	.134	1.826	.181
	Consensus	.755	12.409	.001
Error	Mean Squared Error	5.59E+16		
	Matching	.187		
	Consistency	.073		
	Consensus	.061		

See Panel B for variable descriptions.

Table 7: Supplementary analysis

Panel A: Questions

	<u>Points</u>	<u>Statement</u>
a.	26.30	I looked <i>equally</i> for relationships between spoilage and other costs <i>regardless of the cost pool</i> (that is, <i>cost pools had no effect</i> on my learning strategy) because <i>I did not expect to learn anything about the behavior of the costs from the cost pools.</i>
b.	16.35	I first looked for a relationship between spoilage and the costs in the <i>same cost pool</i> because <i>one person or group of people has responsibility for managing the costs within the pool.</i>
c.	26.78	I first looked for a relationship between spoilage and the costs in the <i>same cost pool</i> because <i>costs within the same pool are likely to have strong relationships.</i>
d.	12.57	I first looked for a relationship between spoilage and the costs in the <i>other cost pool</i> because <i>an important cause of spoilage costs is likely to be included in a different cost pool.</i>
e.	14.62	I <i>ignored all the text descriptions and cost pools</i> and <i>only used the numerical values</i> to learn about relationships between costs.
f.	<u>3.38</u>	Other.
	100.00	Total points

This panel presents follow-up questions. The entries in the points column represent mean responses (this column was blank when presented to participants).

Table 7: Supplementary analysis, continued***Panel B: Comparison of points allocation across cost behavior conditions***

	Location of strong predictor	N	Mean	Std. deviation	Sig. (2-tailed)
Item A	Within pool	40	23.68	25.56	0.34
	Across pools	38	29.05	23.41	
Item B	Within pool	40	17.51	13.95	0.46
	Across pools	38	15.13	14.45	
Item C	Within pool	40	29.31	21.68	0.28
	Across pools	38	24.11	19.78	
Item D	Within pool	40	10.50	14.13	0.17
	Across pools	38	14.74	12.78	
Item E	Within pool	40	15.00	15.67	0.85
	Across pools	38	14.21	21.36	
Item F	Within pool	40	3.75	7.99	0.48
	Across pools	38	2.50	7.51	
Sign					
Item A	Positive	41	28.44	25.31	0.42
	Negative	37	23.92	23.75	
Item B	Positive	41	17.94	15.56	0.30
	Negative	37	14.59	12.38	
Item C	Positive	41	26.30	17.74	0.84
	Negative	37	27.30	24.00	
Item D	Positive	41	13.05	14.99	0.74
	Negative	37	12.03	11.99	
Item E	Positive	41	13.17	18.31	0.47
	Negative	37	16.22	18.91	
Item F	Positive	41	1.10	4.11	0.01
	Negative	37	5.41	9.96	

See Panel A for the post-experimental questions (Items A through F)

Table 7: Supplementary analysis, continued*Panel C: Correlations among points allocations and prediction errors*

Strong Predictor = Within pool (n = 40)							Mean Squared Error
	Item A	Item B	Item C	Item D	Item E	Item F	
Item B	-0.34						
Item C	-0.63	0.07					
Item D	-0.26	-0.15	-0.13				
Item E	-0.15	-0.34	-0.21	-0.14			
Item F	-0.17	0.07	-0.14	-0.02	-0.08		
Mean Squared Error	-0.03	0.11	-0.15	0.19	0.01	-0.05	
Mean Absolute Error	-0.03	0.09	-0.12	0.26	-0.04	-0.12	0.97

Strong Predictor = Across pools (n = 38)							Mean Squared Error
	Item A	Item B	Item C	Item D	Item E	Item F	
Item B	-0.44						
Item C	-0.55	0.24					
Item D	-0.17	-0.24	0.00				
Item E	-0.14	-0.24	-0.48	-0.28			
Item F	-0.17	-0.07	0.04	0.10	-0.21		
Mean Squared Error	-0.30	-0.02	0.32	-0.01	0.08	-0.10	
Mean Absolute Error	-0.26	-0.07	0.28	-0.02	0.12	-0.11	0.98

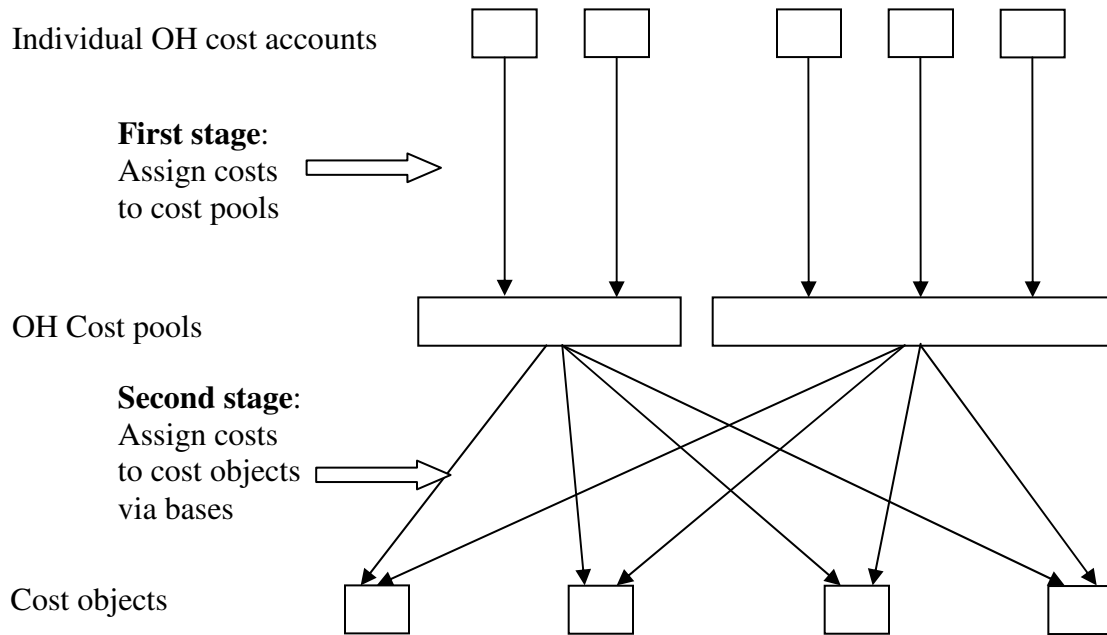
Note: highlighted correlations are significant at $p < 0.10$, two-tailed

See Panel A for the post-experimental questions (Items A through F).

Mean squared error is the mean squared error of participants' predictions compared to predictions derived from the environmental model.

Mean absolute error is the mean absolute error of participants' predictions compared to predictions derived from the environmental model.

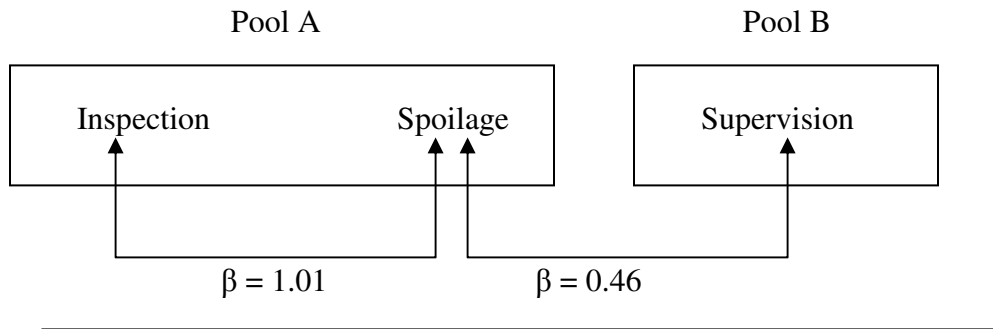
Figure 1: Two-stage overhead cost allocation



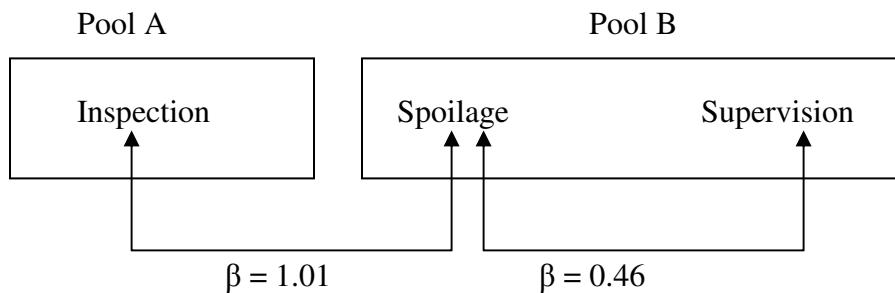
This figure depicts the two-stage process for assigning overhead cost accounts from the chart of accounts to cost objects. In the first stage, overhead costs are assigned to overhead cost pools. In the second stage, overhead costs from the overhead cost pools are assigned to cost objects via allocation bases.

Figure 2: Overhead cost pool classifications

Panel A: Cost pools based on cost allocation



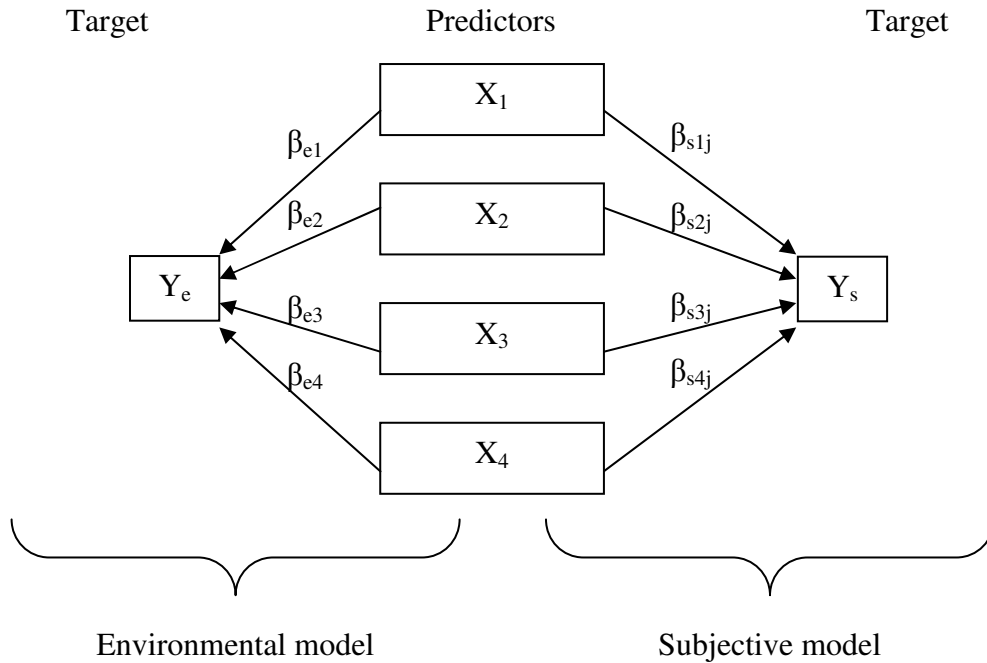
Panel B: Cost pools based on responsibility accounting



This figure depicts two potential overhead cost pool classification decisions. In Panel A, costs are classified such that the stronger relationships (inspection and spoilage) is within a pool, and the weaker relationship (supervision and spoilage) is across pools. This classification might result if the organization decided that allocation of costs to cost objects is the most important use of the cost system.

In Panel B, costs are classified such that the stronger relationship (inspection and spoilage) is across cost pools, and the weaker relationship (supervision and spoilage) is within a pool. This classification might result if the organization decided that responsibility accounting use of the cost system is most important, and spoilage and supervision are the responsibility of the same manager.

Figure 3: Conceptual depiction of lens model method

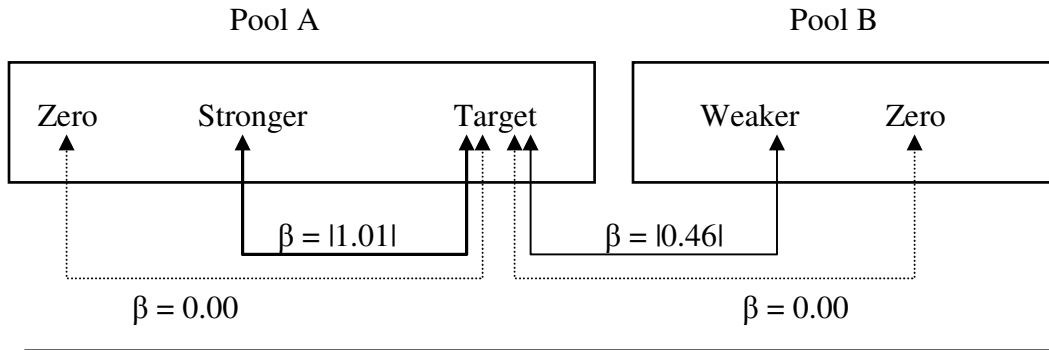


This figure presents a conceptual depiction of the lens model method. The predictors in the center are the four predictors individuals can use to predict the target. The environmental or true model is depicted on the left side of the figure; this model is derived from the data that is presented to individuals and that they could use to estimate the environmental coefficients (β_{ei} , where ei refers to the environmental coefficient on predictor i).

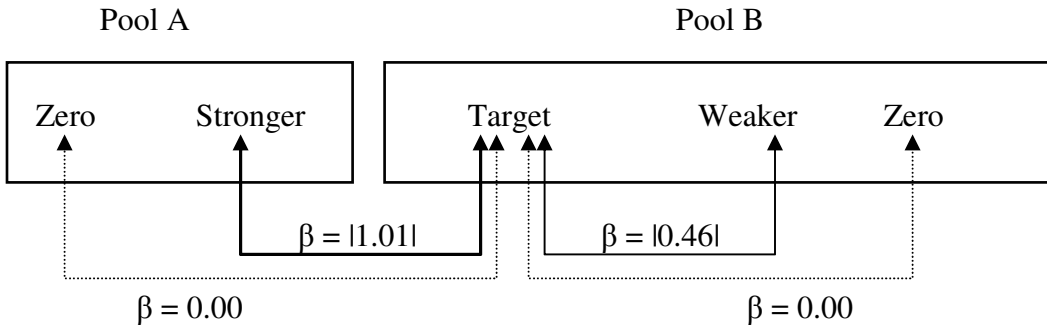
The subjective model on the right side of the figure represents the model of individuals' subjective predictions of the target based on levels of the predictors. Lens model studies regress individuals' predictions of the target on the predictors and estimate each participant's implicit coefficients (β_{sij} , where sij refers to participant j 's implicit coefficient on predictor i).

Figure 4: Overhead cost pool classification and cost behavior

Panel A: stronger relationship is within-pool

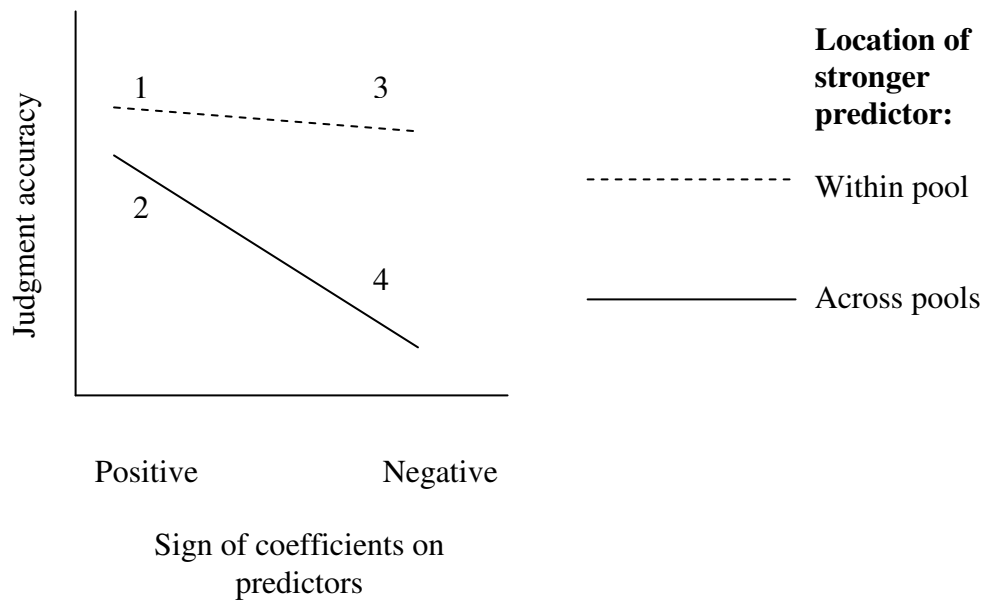


Panel B: stronger relationship is across-pool



This figure presents a conceptual depiction of the design of the experimental task. “Target” describes the target cost that the individual is to predict. “Stronger,” “weaker” and “zero” describe the relationships between the predictors and the target. The β coefficients on the arrows specify the strength of the relationships between the targets and the predictors. The boxes represent overhead cost pool classification boundaries. Note that in this example, all costs have similar scales, so larger coefficients imply stronger correlations and more predictive power. All differences between coefficients are statistically significant at $p < 0.05$.

Figure 5: Graphical presentation of Hypothesis 3



This figure presents the hypothesized interaction between the location of the stronger predictor (either within the same pool as the target cost or across pools) and the sign of the coefficients relating the predictors to the target (positive or negative).