



Mini-review

Testing contingency hypotheses in budgetary research: an evaluation of the use of moderated regression analysis[☆]

Frank G.H. Hartmann^a, Frank Moers^{b, *}

^a*Department of Financial Management, Faculty of Economics and Econometrics, University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands*

^b*Department of Accounting, Faculty of Economics and Business Administration, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands*

Abstract

In the contingency literature on the behavioral and organizational effects of budgeting, use of the Moderated Regression Analysis (MRA) technique is prevalent. This technique is used to test contingency hypotheses that predict interaction effects between budgetary and contextual variables. This paper critically evaluates the application of this technique in budgetary research over the last two decades. The results of the analysis indicate that the use and interpretation of MRA often do not conform to proper methodology and theory. The paper further demonstrates that these problems seriously affect the interpretability and conclusions of individual budgetary research papers, and may also affect the budgetary research paradigm as a whole. © 1999 Elsevier Science Ltd. All rights reserved.

Keywords: Budgetary research; Reliance on accounting performance measures; Budgetary participation; Methodology; Moderated regression analysis; Interaction.

1. Introduction

Over the last 40 years a research paradigm has developed in the management accounting literature that focuses on the use of budgets in organizations. An early study by Argyris (1952) provided a first attempt to describe the effects of using budgets on the behavior of employees. Whereas

Argyris and other researchers in the 1950s and 1960s often studied budget-related issues following a case-study methodology, later studies predominantly relied upon survey data. In the 1970s two such survey studies on budgeting appeared that have become particularly influential. These studies, by Hopwood (1972) and Otley (1978), focused on the behavioral and attitudinal effects of using budgetary information to evaluate the performance of subordinate managers. Hopwood (1972) found that a high reliance on budgetary performance led to a high degree of stress, as well as to dysfunctional managerial behavior. Believing that Hopwood's results were likely contingent on other organizational variables, Otley (1978) designed a study that involved a research site where it was expected that Hopwood's results

* Corresponding author. Fax: +31-20-525-5285.

[☆] The authors gratefully acknowledge the comments made by Ken Merchant, David Otley and two anonymous reviewers. This paper has further benefited from presentations at Maastricht University, the 21st annual EAA meeting, the Fourth International Management Control Systems Research Conference, the EIASM workshop on New Directions in Management Accounting, and the 1999 AAA Management Accounting Research Conference.

would not hold. Indeed, Otley obtained results that were contrary to those of Hopwood. He did not find negative relationships between the use of budgetary performance information and subordinates' attitudes and behaviors; instead he found either no correlation or positive correlations. The conflicting results of these two studies provided an important stimulus to other researchers to adopt contingency perspectives in studying the effects of the formalized construct *Reliance on Accounting Performance Measures* (RAPM). For example, Brownell (1982a) expected that the difference in results could be explained by a construct labeled *Budget Participation* that relates to subordinate managers' involvement in budget setting. This variable had also received ample attention in the literature during the sixties and the seventies (cf. Shields & Shields, 1998). Generally, such contingency studies have aimed to find a match between the use of budgets and the *context* in which they are used. Together, they form a body of literature which has attained a dominant position in contemporary management accounting research (cf. Chapman, 1997). Brownell and Dunk (1991, p. 703) characterize the development of this paradigm as:

The continuing stream of research devoted to this issue constitutes, in our view, the only organized critical mass of empirical work in management accounting at present.

Over the last decade, several papers have provided overviews and evaluations of different aspects of this contingency literature on budgeting (e.g. Briers & Hirst, 1990; Fisher, 1995; Hartmann, in press; Kren & Liao, 1988; Shields & Shields, 1998). The purpose of the present paper is to address and critically evaluate the statistical method used in this literature to test contingency hypotheses. It focuses on the use of Moderated Regression Analysis (MRA), which has become the dominant statistical technique in budgetary research for testing contingency hypotheses. The use of techniques such as MRA has received only little attention in the overview papers mentioned. Yet, such attention seems warranted for at least two reasons. First, since the introduction of the

contingency theory paradigm in budgeting, statistical techniques have become increasingly important (cf. Briers & Hirst, 1990, p. 385). They not only affect the design, execution and success of individual studies, but also determine the paradigm's overall success (cf. Lindsay, 1995). Second, attention to the MRA technique seems particularly warranted given the complexity and specificity of the technique, and the problems associated with its use (e.g. Arnold, 1982, 1984; Jaccard, Turrisi & Wan, 1990). As will be shown in detail below, budgetary papers often appear to neglect these complexities, which causes flaws in the application of MRA and in the interpretation of results.

The remainder of the paper is structured as follows. Section 2 begins with a short explanation of the concept of 'fit' in contingency theory. It continues with an explanation of the basic properties of MRA. Section 3 discusses the selection of budgetary research papers for analysis. Section 4 describes specific characteristics of MRA and presents the findings of the analysis of the use of MRA in the selected research papers. Finally, Section 5 discusses the implications of the findings for both the current state and required future developments of budgetary research.

2. Testing contingency theories of budgeting

Contingency theories of accounting are the opposites of universal theories of accounting in that they link the effects or the optimality of accounting systems to the environment and context in which these systems operate. In a summary of early management accounting studies that used contingency frameworks, Otley (1980) concluded that much needed to be done in the development of a contingency theory of accounting, and he outlined some minimal requirements for such a theory, stating that:

(...) a contingency theory must identify specific aspects of an accounting system which are associated with certain defined circumstances and demonstrate an appropriate matching (Otley, 1980, p. 413).

Three elements in this prescription are essential, relating to: (1) the ‘specific aspects’ (2) the ‘defined circumstances’; and (3) the ‘appropriate matching’. The first element (i.e. specific aspects) points to the demand for specificity of the accounting system variables in the formulation and test of theories. The second element (i.e. defined circumstances) points to the *conceptual* difference between a universal theory and a contingency theory. The third and last element (i.e. appropriate matching) forms the core of this paper, as it points to the *empirical* difference between a universal theory and a contingency theory. Otley (1980) does not show how an ‘appropriate matching’ is to be defined theoretically, nor does he prescribe how it is to be determined empirically. Such prescriptions and illustrations *can* be found, however, in the organizational literature, which has a larger history in contingency methodology (cf. Chapman, 1997) and pays ample attention to the theoretical and empirical aspects of determining the ‘appropriate matching’. In the remainder of this paper, this ‘appropriate matching’ element will be addressed with the more common term ‘contingency fit’.

An overview of the organizational literature reveals several different concepts of ‘contingency fit’ (see, e.g. Drazin & Van de Ven, 1985; Schoonhoven, 1981; Venkatraman, 1989). These concepts of fit are each associated with a different theoretical interpretation, and each require a different statistical test. In this paper, the discussion is restricted to a type of contingency fit called the interaction type of fit, which is the dominant conceptualization of contingency fit in budgetary research. The appropriate statistical technique to test the interaction type of fit is through Moderated Regression Analysis (MRA), that will be explained below. Appendix A briefly outlines several other common types of fit, stating both the typical format of the underlying contingency hypothesis and the appropriate statistical test.

2.1. Moderated Regression Analysis, the basic format

Moderated Regression Analysis (MRA) is a specific application of multiple linear regression analysis, in which the regression equation contains

an ‘interaction term’ (e.g. Champoux & Peters, 1987; Southwood, 1978). A typical equation for the multiple regression of a dependent variable (Y) on two independent variables (X_1 and X_2) is presented in Eq. (1):

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon \quad (1)$$

In contrast, a typical regression equation used in MRA has the format of Eq. (2):

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 \times X_2 + \varepsilon \quad (2)$$

Eq. (2) differs from Eq. (1) by the inclusion of the product of the two independent variables ($X_1 \times X_2$). This product term is said to represent the *moderating effect* of variable X_2 on the relationship between X_1 and Y . In contrast, the other terms in the equation (X_1 and X_2) are said to represent the *main effects* of variables X_1 and X_2 on Y . The meaning of this product term in establishing a moderating effect can be illustrated by taking the partial derivative of Eq. (2) with respect to X_1 ($\partial Y / \partial X_1$), which has the format expressed by Eq. (3):

$$\partial Y / \partial X_1 = \beta_1 + \beta_3 X_2 \quad (3)$$

As Eq. (3) illustrates, the term representing the partial derivative ($\partial Y / \partial X_1$) is a function of X_2 . This means that the ‘form’ of the relationship between Y and X_1 is a function of X_2 , or in short, that variable X_2 *moderates* the form of the relationship between X_1 and Y (cf. Champoux & Peters, 1987, p. 244; Jaccard et al., 1990, p. 22). A moderating effect can be graphically illustrated as a variation in the slope of the regression line of Y and X_1 as a function of X_2 . Fig. 1 below depicts a situation in which the slope of the regression line between X_1 and Y is more positive for higher values of X_2 . This is expressed by stating that Y is a function of the *interaction* between X_1 and X_2 . Alternatively, it is said that the relationship between Y and X_1 is *contingent* upon X_2 . To prove the *contingency* hypothesis, therefore, one must prove that X_2 influences the relationship between X_1 and Y , or that X_1 and X_2 *interact* to affect Y .

Although X_2 is considered the moderator in the example above, a similar analysis applies if X_1 is

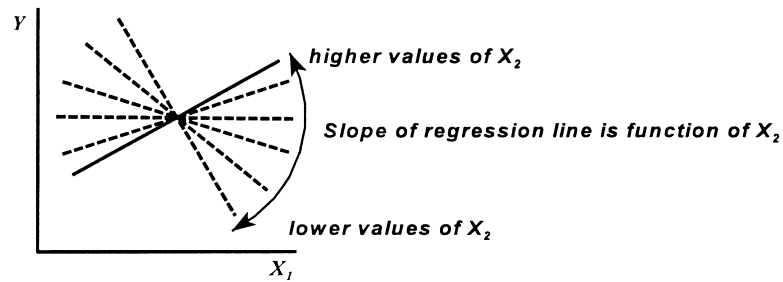


Fig. 1. Moderating effect.

considered the moderator of the relationship between X_2 and Y . Then, illustrated with Eq. (3a), the partial derivative is taken with respect to X_2 ($\partial Y/\partial X_2$), and it follows that the relationship between X_2 and Y is a function of X_1 .

$$\partial Y/\partial X_2 = \beta_2 + \beta_3 X_1 \quad (3a)$$

For this reason, the moderating effect expressed by the interaction term in Eq. (2) is called 'symmetrical' (cf. Arnold, 1982; Southwood, 1978). This implies that if X_2 moderates the relationship between X_1 and Y , then X_1 necessarily also moderates the relationship between X_2 and Y . It is because of this symmetry that the neutral expression refers to an 'interaction' of X_1 and X_2 to affect Y . Whether an independent variable is labeled as a moderator or an independent variable is a matter of theory rather than statistics. A moderator variable theoretically affects the relationship between the independent variable and the dependent variable, but is not itself theoretically related with either the dependent or independent variables (e.g. Arnold, 1982, p. 154; 1984, p. 216; Shields & Shields, 1998, p. 51). Typically, variables are labeled moderators that are *exogenous* or *uncontrollable* (e.g. Cohen & Cohen, 1983, p. 305). Sharma, Durand, and Gur-Arie (1981) provide an overview and taxonomy of moderator variables.

In empirical contingency research, MRA is used to establish the existence of a statistically significant interaction effect. A method to do so is through hierarchical regression analysis (e.g. Arnold & Evans, 1979; Cohen & Cohen, 1983; Cronbach, 1987; Southwood, 1978). This method requires running two regressions, one with the

main-effects-only [cf. Eq. (1)] and a second with both main effects and the interaction term [cf. Eq. (2)]. A significant interaction effect is confirmed by the statistical significance of the additional variance explained by the inclusion of the interaction term (i.e. the significance of the increase in R^2). This method is equivalent to the simpler and more direct assessment of the significance of the t -value associated with the coefficient of the product term (see Southwood, 1978, p. 1168; Arnold, 1982, p. 157; Jaccard et al., 1990, p. 22). The equivalence of these two methods is illustrated by Cohen and Cohen (1983), who show that the F -statistic for the increase in R^2 equals the square of the t -statistic for the interaction term.¹

The equivalence of these two methods is illustrated by Cohen and Cohen (1983), who show that the F -statistic for the increase in R^2 equals the square of the t -statistic for the interaction term. In the example used above, this means that a test for a statistically significant moderating effect of X_2 on the relationship between X_1 and Y implies a test whether the coefficient of the interaction term (β_3) in Eq. (2) is statistically significant. The symmetry applies here as well; a significant t -value of the coefficient of the interaction term thus simultaneously implies a significant moderating effect of X_1 on the relationship between X_2 and Y .

¹ See, for example, Chenhall (1986) for a redundant test for the significance in incremental explanatory power after testing for the significance of the interaction term. Surprisingly, the F -statistic of incremental explanatory power and the squared t -statistic of the interaction in this study do not match (F equals 7.27, t -square equals 28.30). The only explanation seems to be a calculation error.

The interaction in Eq. (2) above is commonly called a *two-way interaction*, since the equation contains two variables and their interaction. Moreover, given that in the example (see Fig. 1) the relationship between X_1 and Y is more positive (or: less negative) for higher values of X_2 , it is called a ‘positive interaction’ between X_1 and X_2 . A ‘negative interaction’ signifies that the relationship between X_1 and Y is *more negative* (or: less positive) for higher values of X_2 . Additionally, the interaction can be either *monotonic* or *non-monotonic*. A monotonic interaction exists when the partial derivative does not ‘cross’ the horizontal axis. This means that the moderating effect of X_2 changes the slope of the relationship between X_1 and Y within positive values, or negative values, only (cf. Schoonhoven, 1981). Appendix A presents verbal examples of both monotonic and non-monotonic interactions.

2.2. Moderated Regression Analysis with a dummy variable

A special and often used form of MRA is obtained when the moderator variable is a dummy variable, taking on only discrete values (e.g. 0 and 1). If, for the example above, the moderator X_2 has only values of 0 and 1, the original equation expressing the interaction effect between X_1 and X_2 in Eq. (2), can be rewritten in the formats of Eq. (2a) and Eq. (2b) below.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 \times X_2 + \varepsilon \quad (2)$$

$$Y = \beta_0 + \beta_1 X_1 + \varepsilon \quad (\text{for } X_2 = 0) \quad (2a)$$

$$Y = (\beta_0 + \beta_2) + (\beta_1 + \beta_3)X_1 + \varepsilon \quad (\text{for } X_2 = 1) \quad (2b)$$

While this does not change the interpretation of the coefficient of the interaction term (β_3), the decomposition illustrates that an analysis is done for two subgroups. Therefore, the MRA with a dummy variable is sometimes called ‘subgroup regression analysis’ (e.g. Stone & Hollenbeck, 1984) in which the ‘subgroups’ are distinguished based on extreme (e.g. high and low) values of the moderator variable. Arnold (1984, pp. 219–221) argues that the label ‘subgroup regression analysis’ may lead to the incorrect belief that this method differs from general MRA. In fact, MRA is always concerned with different ‘subgroups’, however, introducing a dummy variable reduces the number of ‘subgroups’ to two. A graphical example of MRA when the moderator variable is a dummy variable is presented in Fig. 2.

Fig. 2 shows two regression lines, one for each of the two values of X_2 . The figure illustrates a positive interaction, which means that the coefficient (β_3) of the interaction term is positive. From comparing Eqs. (2a) and (2b) above, it follows that a positive and significant coefficient (β_3) suggests that the slope of the regression line for the ‘ $X_2 = 1$ ’ subgroup is significantly ‘more positive’ than the slope of the regression line for the ‘ $X_2 = 0$ ’ subgroup. It should be noted that the associated label ‘positive interaction’ is only meaningful if the dummy values 0 and 1 reflect underlying values (e.g. low and high) of the moderator variable. However, MRA is also meaningfully applied if a dummy does not reflect an underlying quantitative variable, for example, if it reflects a natural dummy

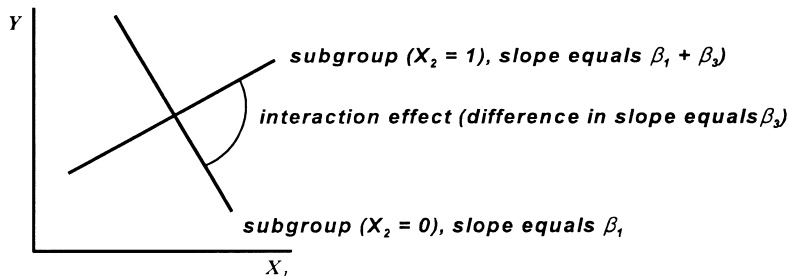


Fig. 2. Interaction effect when the moderator is a dummy variable.

(e.g. male–female). Despite the prevalence of such ‘natural dummies’, subgroup regression analysis is commonly performed based on a categorization of the scores on an underlying continuous variable. Such categorization has been argued to be unadvisable, since it implies a loss of information (e.g. Cohen & Cohen, 1983, p. 310; Pedhazur & Pedhazur-Schmelkin, 1991, p. 539), yet it has substantial advantages relating to the understandability of the MRA outcomes and the statistical power of the MRA technique (Arnold, 1984, pp. 221–222). These advantages are especially important when the analysis incorporates interactions of a higher order than the two-way interactions discussed so far. Such higher-order interactions are discussed further below.

3. Selection of the budgetary research papers

The budgetary research papers were selected for analysis using four criteria. These were: (1) research method; (2) publication date; (3) journal of publication; and (4) subject of study. The first criterion aimed at the selection of papers that used a survey methodology using questionnaires. The second criterion resulted in papers published in the period from 1980 to 1998. This period was chosen because it lies after the influential conceptual paper of Otley (1980). The third criterion aimed at selecting papers from high-quality accounting journals. This criterion was applied to limit the total number of papers, as well as to exclude ‘lower journal quality’ as a potential explanation for the findings. Based on an examination by Brown and Huefner (1994) of the perceived quality of accounting journals among different respondents, papers were selected from *The Accounting Review*, *Journal of Accounting Research*, and *Accounting, Organizations and Society*.² Finally, regarding the subject of study, papers were selected that hypothesize and test contingency fit concerning budget-related variables, such as *RAPM* and

Budgetary Participation, using an ‘interaction’ concept of fit. The application of these criteria resulted in the selection of 28 papers. Table 1 shows the dispersion of the papers over the three journals. Appendix B provides an overview of the 28 papers that meet the selection criteria.

4. Analysis of the budgetary research papers

This section analyzes the application and interpretation of MRA in the 28 budgetary research papers reviewed. In addition to the basic format of MRA discussed above, here different subsections explain different specific characteristics of MRA. These specific characteristics relate to interaction and (1) the strength of relationships; (2) the formulation of hypotheses; (3) lower-order effects; (4) multiple and higher-order interactions; (5) effect size; and (6) (non-) monotonicity. Each subsection illustrates and discusses how these MRA characteristics appear in the individual studies.

4.1. Interaction and the strength of relationships

In Section 2.2 above, a reference was made to the ‘subgroup’ method for illustrating the difference in the slope of the regression line for two subgroups. The literature uncovers another use of subgroup analysis which tests for differences between subgroups in the *strength* of the relationships between the independent and dependent variables (e.g. Champoux & Peters, 1987, p. 243; Stone & Hollenbeck, 1984). Within the context of the example, this ‘subgroup correlation analysis’ implies that a test is made for differences between the correlation of X_1 and Y for extreme values of

Table 1
Dispersion of reviewed budgetary articles over journals

Journal (acronym used)	Number of articles
<i>The Accounting Review</i> (TAR)	5
<i>Journal of Accounting Research</i> (JAR)	6
<i>Accounting, Organizations and Society</i> (AOS)	17
Total	28

² The *Journal of Accounting and Economics* belongs to the four journals with the highest perceived quality in the Brown and Huefner (1994) study. It has not published, however, contingency studies of budgeting.

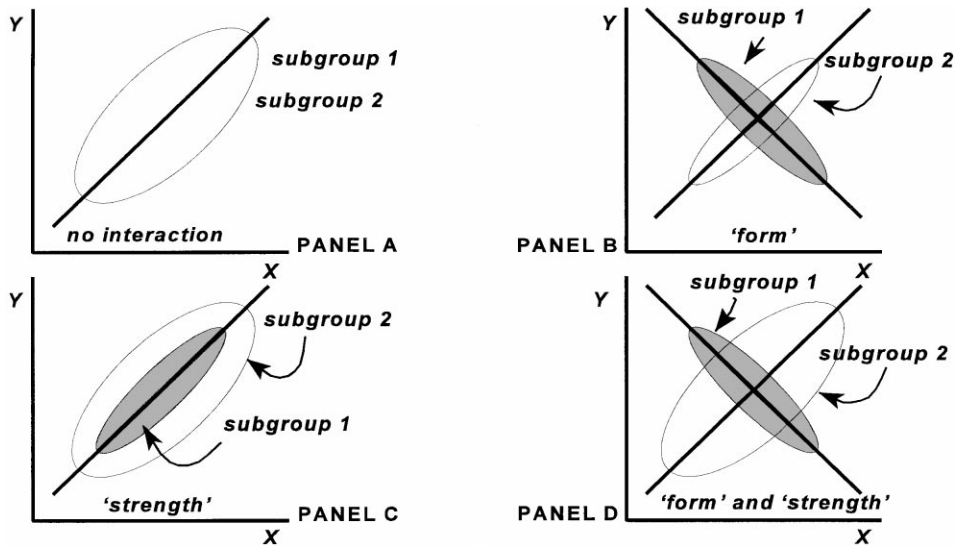


Fig. 3. Interaction as strength and as form.

X_2 . The difference in substantive meaning of the two kinds of analyses is graphically illustrated in Fig. 3. In this figure, the shaded areas represent the ‘clouds’ of observations (scatter diagrams) of the relationship between X and Y . A ‘wide’ cloud indicates a low correlation and a ‘narrow’ cloud indicates a high correlation. For each cloud the appropriate regression lines are depicted as well.

Panel A of Fig. 3 shows no sign of interaction, since both correlations and regression lines are equal for the two subgroups. Panel B illustrates a ‘form’ interaction since the slope of the regression line is different for the two subgroups. No indication of interaction for ‘strength’ exists, since X appears to predict Y in both subgroups equally well, evidenced by the absolute values of the correlation coefficients being equal. Panel C is the opposite of panel B since there is no difference in slope, but there is a difference in ‘strength’. This means that in subgroup 1, X is a better predictor of Y than in subgroup 2. Panel D shows a combination of ‘form’ and ‘strength’ interactions, since both the slope of the regression line and the correlation of X and Y are different for the two subgroups.

There has been some confusion in the literature about whether contingency hypotheses reflect differences in ‘form’ or in ‘strength’ (see, e.g. Arnold,

1982, 1984; Stone & Hollenbeck, 1984). The common understanding, however, now seems to be that contingency hypotheses are of the form type (panels B and D, Fig. 3), and it is even argued that differences in strength are commonly meaningless (Arnold, 1982, pp. 153–154; Schmidt & Hunter, 1978, p. 216). A problem, therefore, exists with respect to the relationships hypothesized and tested in many budgetary studies. In 15 of the 28 papers used in this paper there is either no hypothesis explicitly stated (Brownell & Dunk, 1991; Brownell & Hirst, 1986; Brownell & Merchant, 1990) or it is stated in a null form predicting that there is ‘no interaction’ between the measured variables (Brownell, 1982a, b, 1983, 1985; Chenhall, 1986; Dunk, 1989, 1990, 1993; Frucot & Shearon, 1991; Harrison, 1992, 1993; Mia & Chenhall, 1994). This raises the question about the meaning of the word ‘interaction’ in these cases, in particular whether it relates to the *strength* of the relationship or the *form* of the relationship. Since these types of fit are not equivalent for either the theoretical interpretation or the statistical test, null hypotheses are inadequate for describing the specific contingency formulations and statistical test to be used. Obviously, if an author states that there is an ‘interaction’ (e.g. Hirst & Lowy, 1990),

this hypothesis has the same shortcomings as the null hypothesis described above.

Despite the dominance of form interactions in contingency research, budgetary researchers *have* investigated hypotheses that seem to express *strength* interactions. If there is a theory supporting such relationships, the appropriate statistical analysis consists of calculating *z*-scores of the correlation coefficients. Since the interest is in the difference in predictive power between subgroups, it is important that the *z*-scores are calculated using absolute values of the correlation coefficients. Obviously, their sign does not contain information about predictive power. Merchant (1981, 1984, 1990) and Govindarajan (1984) test for differences in correlation coefficients, but do not calculate nor specifically mention the ‘absolute’ *z*-score.

4.2. Interaction and the formulation of hypotheses

The examples above point to a weak link between the verbal and statistical format of hypotheses, which has been noted and discussed before by Schoonhoven (1981). She criticizes studies in the organizational literature for their lack of clarity in stating contingency hypotheses, which affects the obviousness of the statistical test to be used. Above, a reference was made to Govindarajan (1984), who tests a strength hypothesis. A further problem with Govindarajan’s (1984) analysis relates to the substantive content of his contingency hypothesis, which is incompatible with his theoretical arguments.³ His hypothesis predicts that ‘organizational effectiveness’ affects the strength of the relationship between ‘environmental uncertainty’ and ‘evaluation style’. The format of the statistical analysis is in conformity with this hypothesis, since it compares the correlation coefficients for the two subgroups. The subsequent interpretation of these results, however, is not. Govindarajan (1984, p. 133) concludes that environmental uncertainty has a ‘significant

moderating effect’ on the relationship between evaluation style and organizational effectiveness. This is a peculiar interpretation since the subgroup correlation analysis is based on high and low ‘effectiveness’ subgroups. This suggests that ‘effectiveness’ is the moderator instead of the theoretically relevant moderator ‘environmental uncertainty’.⁴ Overall, therefore, there is no match between the theoretical arguments, the formulation of the hypothesis, and the interpretation of the statistical analysis of the hypothesis. Govindarajan’s (1984) conclusion about the moderating effect of ‘environmental uncertainty’ is clearly unfounded.

In another six of the 28 papers (Merchant, 1981; 1984; Brownell, 1982b; Imoisili, 1989; Frucot & Shearon, 1991; Dunk, 1992), obvious differences exist between the hypothesis and the statistical test used, which makes the conclusions by the authors debatable.⁵ Merchant (1981, 1984) hypothesizes an interaction affecting the form of the relationship, as shown by an example of the following hypothesis:

Organizational performance tends to be *higher* where there is a ‘fit’ between the use of budgeting and the situational factors, as described in Hypotheses 1–4. (Merchant, 1984, p. 294; emphasis added.)⁶

Recall from the previous section that the statistical test used is a correlation per subgroup and relates to the strength of the relationship. The results of the analysis are therefore of little relevance and value to the contingency hypothesis stated.

A second example of a difference between the hypothetical and measured fit are the papers of

³ Govindarajan’s (1984) theoretical arguments indicate an interaction of the form type. Govindarajan even *specifically* proposes a non-monotonic relationship (see Fig. 1, p.128)

⁴ Note that labeling an independent variable the ‘moderator’ in MRA is a question of theory because of the symmetry mentioned earlier. Yet, Govindarajan does not switch the moderator and the independent variable, but switches the moderator and the *dependent* variable. In this case, the symmetry does not apply.

⁵ The papers of Imoisili (1989) and Dunk (1992) will be discussed in Sections 4.4 and 4.5, respectively.

⁶ This hypothesis is the fifth in Merchant’s paper and refers to hypotheses 1–4. In these hypotheses, Merchant applies the selection type of fit and uses correlation analysis. For example hypothesis 4 states: “Larger, more diverse departments tend to place greater emphasis on formal budgeting.”

Brownell (1982b) and Frucot and Shearon (1991, p. 85). Frucot and Shearon (1991) state a null hypothesis in the following form:

In Mexico, locus of control does not have, through an interaction effect with budgetary participation, a significant effect on performance.

Subsequently the following equation is statistically tested:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 - X_2 + \varepsilon \quad (4)$$

This equation does not, however, measure fit as interaction but fit as matching (Venkatraman, 1989, p. 431). Because fit as interaction (i.e. form) is hypothesized, the verbal statement and statistical measure are not compatible.⁷ Given this incompatibility, the conclusion cannot be interpreted and the statistical results become exploratory at best. The reason the authors give for using the matching variable is that:

(...) the multiplicative interaction term does not provide a good measure of this matching condition” (Frucot & Shearon, 1991, p. 90).

Although this is true, the hypothesis expresses ‘fit’ as interaction and simply does not predict the matching condition. Frucot and Shearon (1991, p. 90) also state that the variable should be measured by the absolute difference, arguing that this better reflects the relationship among the three variables. If so, the hypothesis should have been altered to state the matching condition, and the alternative and only proper formulation would have been:⁸

In Mexico, given the value of locus of control, there is a unique value of budgetary participation that produces the highest value of performance; deviations from this relationship

in either direction reduce the value of performance. (cf. Schoonhoven, 1981).⁹

4.3. Interaction and lower-order effects

As illustrated above, the equation representing an interaction effect [Eq. (2)] includes not only the interaction term ($X_1 \times X_2$), but also the two main effects (X_1) and (X_2). At least three topics related to the inclusion of lower-order effects warrant discussion. These are (1) the reason for including lower-order effects; (2) the interpretation of the coefficients for lower-order effects found in the regression analysis; and (3) the potential problem of multicollinearity. First, the reason for the inclusion of lower-order effects in MRA is to prevent conclusions of the existence of an interaction effect when such an effect is solely due to lower-order effects. An example is found with Hirst (1983) who tests a model represented by Eq. (5) below that does not include main effects:

$$Y = \beta_0 + \beta_1 X_1 \times X_2 + \varepsilon \quad (5)$$

In this case, finding a significant coefficient β_1 does not necessarily indicate the existence of an interaction effect, since it could be due to a significant relationship between X_1 and Y regardless of X_2 . In other words, because the interaction term is the product of the two main effects, it is likely to ‘steal variance’ from its constituting parts (cf. Cohen & Cohen, 1983, p. 305). Stone and Hollenbeck (1984, p. 201) argue in this respect:

(...) while the cross-product term in a regression equation “carries” the interaction, the same cross-product term is not the interaction.

This implies that when testing for an interaction effect, the lower-order effects should be ‘partialled out’ by including them in the regression equation (Southwood, 1978, p. 1164; Cohen & Cohen,

⁷ The authors also state that the relationship is ‘accentuated’ and ‘attenuated’ (p. 85). This refers to accelerating and decelerating effects and thus to the form of the relationship.

⁸ The same line of reasoning applies to Brownell (1982b).

⁹ For a discussion of the use of difference scores, see Bedeian, Day, Edwards, Tisak, and Smith (1994).

1983, p. 348; Stone & Hollenbeck, 1984, p. 201). Consequently, Hirst's (1983) conclusion based on the results of Eq. (5) is unfounded.¹⁰

A second issue relates to the proper interpretation of coefficients obtained for lower-order effects in MRA. Southwood (1978, p. 1168) argues that such coefficients generally have no theoretical meaning. The reason is that in the behavioral sciences the variables are usually measured using interval scales, and not ratio scales. This means that scale origins and thus linear transformations of variable scores are arbitrary, and have no substantive meaning. Southwood (1978, p. 1668, pp. 1198–1201) shows that such linear transformations *do* change the coefficient of the lower-order effects in the MRA equations, and therefore, these coefficients cannot be easily interpreted. This does not imply that the coefficients of the lower-order effects in an interaction model are lacking *all* meaning. In particular, they signify the main effect of the variable (e.g. X_1) when the value of the other variable (X_2) is zero (Jaccard et al., 1990). Only for ratio scale variables is this zero 'meaningful' (Southwood, 1978, p. 1165). For interval scale variables, the zero value obtains a specific

meaning if the variables are 'centered' around their respective means. In this case, the coefficients of the main effects represent the effect of one variable at the (sample) average of the other (Jaccard et al., 1990, p. 34). Despite this meaning, it is important to point out that the coefficients obtained for the main effects when applying MRA are, in general, different from those that would be obtained through a regression model without the interaction term. Furthermore, it is important to note that linear transformations do not change the coefficient of the interaction term, nor its *t*-statistic and level of significance (Cohen & Cohen, 1983, pp. 305–306; Southwood, 1978, p. 1168). Thus, Brownell (1982a), Chenhall (1986) and Mia (1988) are wrong in assuming that 'centering' the variables provides a:

clearer basis for predicting *the sign of* (...) the coefficient of the interaction term (Brownell, 1982a, p. 20; emphasis added).

Since the coefficient of the interaction term is not sensitive to the scale origins, it also follows that for tests of an interaction effect using MRA the independent variables need not be ratio-scaled (Southwood, 1978, p. 1167; Arnold & Evans, 1979).

The analysis of the budgetary papers shows that the interpretation of lower-order effects is subject to both Type-I and Type-II errors; interpretation of *invalid* lower-order effects and no interpretation of *valid* lower-order effects. First, five papers show an invalid interpretation of main effects in two-way interactions (Brownell, 1982a, 1983, 1985; Chenhall, 1986; Mia, 1989). For example, Brownell (1982a) studies (among others) the moderating effect of *Budget Emphasis* on the relationship between *Budgetary Participation* and *Performance*. He interprets the interaction effect, and both main effects [see Eq. (2)]. The *Budgetary Participation* variable in the MRA equation is measured as a 'deviation score' from the overall mean (i.e. centered). A linear transformation of the raw score of one independent variable leads to a change in the regression coefficient of the other independent variable. This implies that the coefficient of the main effect of *Budget Emphasis* changes. As a result, the coefficient of *Budget Emphasis* illustrates

¹⁰ Hirst (1983, p. 600) claims to have found an interaction effect of the form-type based on subgroup regression-analysis. Since the regression coefficients are standardized, and since standardized regression coefficients in simple linear regression equals correlation coefficients, the subgroup analysis is in fact subgroup correlation-analysis. This latter statistical format, however, tests *strength* and not *form*, so that Hirst's findings seem unfounded. However, in this specific case additional information suggests the existence of a *form* interaction for one of the two dependent variables (job-related tension). The correlation coefficients between RAPM and job-related tension for low and high task uncertainty are -0.33 and 0.53 respectively. Both coefficients are statistically significant, which means that the regression coefficients per subgroup are also significant. As a result, there will be a significant *negative* coefficient for the low task uncertainty subgroup and a significant *positive* coefficient for the high task uncertainty subgroup. Since significant means significantly different from zero, a positive (negative) significant coefficient is also significantly different from any negative (positive) coefficient. Thus, the difference between these two regression coefficients is also significant. The significance of this difference is equal to the significance of the interaction term in MRA and the fit as interaction (form) is supported. Note that Hirst's (1983) paper, as it is published, thus does *not* prove a fit as interaction (form); the above analysis does.

the effect of *Budget Emphasis* on *Performance* at the average level of *Budgetary Participation*.¹¹ This interpretation differs from Brownell's (1982a, pp. 20–21) who incorrectly interprets the regression coefficient unconditional on *Budgetary Participation*. Further, the *Budget Emphasis* variable is *not* centered around its mean, which makes the coefficient of *Budgetary Participation* uninterpretable. Brownell's conclusion regarding *Budgetary Participation* is therefore unfounded.¹² In studies using higher-order interactions, similar problems are found (e.g. Brownell & Hirst, 1986; Imoisili, 1989). These studies are discussed below in the section on higher-order interactions.

The above five studies precede the introduction of the Southwood (1978) paper into the management accounting research literature.¹³ The more recent and uncritical use of Southwood's paper, however, has led to the reverse situation; no interpretation of valid lower-order effects. Examples include the following. Mia and Chenhall (1994) study the moderating effect of *Function* on the relationship between the *Use of Broad Scope MAS* and *Managerial Performance*. The *MAS* variable is a 'difference score' from the overall mean (i.e. centered). This means that the coefficient of the main effect of *Function* is interpretable. The authors refer to Southwood and state:

No attempt was made to interpret the coefficients (...) that related to extent of use of broad scope Management Accounting System (MAS) information or function. (p. 9; emphasis added.)

Thus, they incorrectly state that the coefficient of *Function* cannot be interpreted and therefore lose information. Using this information would have

allowed the supplementary conclusion that *Function* by itself has a direct influence on managerial performance, in that marketing managers perform better than production managers at the 'average' *MAS*. In Lau, Low and Eggleton (1995) all independent variables are centered. This means that the lower-order effects in their interaction models could have been interpreted. The authors, however, refer to Southwood and reject the interpretation of lower-order effects (p. 369). Although, the papers of Mia and Chenhall (1994) and Lau et al. (1995) are the only concrete and explicit examples of a Type-II error, seven other studies specifically refer to Southwood and a priori reject the interpretation of lower-order effects (Brownell & Dunk, 1991; Dunk, 1990, 1992, 1993; Harrison, 1992, 1993; Gul & Chia, 1994). The conclusion seems warranted that researchers are unaware of the possibilities and problems of interpreting lower-order effects in MRA. In a way, Southwood's paper has had both a positive and a negative effect on this. The positive effect is that main effects indeed cannot usually be interpreted, while the negative effect has been that it is believed that they should never be interpreted in MRA. Unfortunately, this false idea even shows up in handbooks aimed at novice management accounting researchers (Brownell, 1995). Here it says that:

It is now widely understood that the estimated coefficients for variables, included in an equation along with their cross product with another variable, are *not* interpretable. (Brownell, 1995, p. 55; emphasis added.)

It also says that Southwood has shown that if a constant is added to interval scale data that:

(...) this will alter the estimated coefficient for the variable to which the constant was added (p. 55; emphasis added).

This is, of course, incorrect. The coefficient of the variable to which the constant is *not* added, changes. The 'easy' reference to papers like Southwood's, without critically evaluating its wisdom, is remarkable at least.

¹¹ More specifically, the results show that at the average level of *Budgetary Participation*, lower *Budget Emphasis* leads to increased *Performance*.

¹² Brownell concludes that higher *Budgetary Participation* is associated with higher *Performance*.

¹³ Schoonhoven (1981) explicitly incorporates several main effects in her multiple interaction equation and incorrectly interprets them all. This is peculiar since Southwood (1978) is among her references.

The third and final point related to lower-order effects in MRA is the potential problem of multicollinearity, caused by the fact that the lower-order effects (e.g. X_1 and X_2) and their product are likely to be correlated (e.g. Drazin & Van de Ven, 1985). Yet, from the same arguments that show the insensitivity of the coefficient of the interaction term in MRA ($X_1 \times X_2$) to changes in scale origins of X_1 and X_2 , it follows that multicollinearity is not a problem when applying MRA (cf. Dunlap & Kemery, 1987). In particular, it is possible to use linear transformations of X_1 and X_2 that remove the correlation between the main terms and the interaction term (Southwood, 1978, p. 1167; Jaccard et al., 1990, p. 22). The ‘correct’ linear transformation of variable scores to minimize the correlation between the independent variables and their product is the centering procedure discussed before (e.g. Jaccard et al., 1990, p. 34). Recall that since no such linear transformation ever affects the coefficient of the interaction term, the coefficient of the interaction term is always interpretable, and centering is not required. As was illustrated, this is not always recognized in budgetary studies (Brownell, 1982a, 1983; Chenhall, 1986; Mia, 1988, 1989; Lau et al., 1995). As an example, Lau et al. (1995) use deviation scores from the mean because, as they state, it reduces the ‘problem’ of multicollinearity (p. 368). However, no such problem exists.¹⁴

4.4. Multiple and higher-order interactions

The previous sections discussed the general format of MRA and focused on the analysis of a single two-way interaction. MRA, however, is not restricted to the analysis of a single two-way interaction. It can also be used to analyze (1) multiple two-way interactions, and (2) any n -way interaction. First, multiple two-way interactions can be tested by extending Eq. (2) to include an additional main and interaction effect, as shown by the example of Eq. (6a):

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_1 \times X_2 + \beta_5 X_1 \times X_3 + \varepsilon \quad (6a)$$

As for any equation, the inclusion of an additional two-way interaction in MRA should be based on theory (Jaccard et al., 1990, p.40–41). More specifically, the theory should state that the relationship between X_1 and Y is not only a function of X_2 , but also of X_3 . In other words,

$$\partial Y / \partial X_1 = \beta_1 + \beta_4 X_2 + \beta_5 X_3 \quad (6b)$$

Imoisili (1989, p. 327) exactly hypothesizes the relationship depicted by Eqs. (6a) and (6b). However, this hypothesis is incorrectly tested, since the equation analyzed also includes a redundant two-way interaction term $X_2 \times X_3$, and a redundant three-way interaction term $X_1 \times X_2 \times X_3$. The result of analyzing this equation does not provide an answer to the hypothesis, since the partial derivative ($\partial Y / \partial X_1$) of the function that Imoisili uses is given by Eq. (7), which contains an ‘interaction effect’ ($X_2 \times X_3$) not hypothesized.

$$\partial Y / \partial X_1 = \beta_1 + \beta_4 X_2 + \beta_5 X_3 + \beta_7 X_2 \times X_3 \quad (7)$$

The only correct statistical test would have been Eq. (6a). Harrison (1992, 1993) provides an example of similar problems. Harrison (1992) uses Eq. (8) to test for the existence of two-way interactions between *Budgetary Participation* (X_1), *RAPM* (X_2), and the dummy variable *Nation* (X_3) to affect *Job-Related Tension* (Y):

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_1 \times X_2 + \beta_5 X_1 \times X_3 + \beta_6 X_2 \times X_3 + \varepsilon \quad (8)$$

Harrison (1992, p. 11) is only interested in the moderating effect of X_1 on the relationship between X_2 and Y , and his analysis is therefore flawed for two reasons.¹⁵ First, there is no theoretical foundation for using the multiple interaction equation. Only the first interaction ($X_1 \times X_2$) is hypothesized by Harrison, both other interactions

¹⁴ The only ‘problem’ that multicollinearity can cause is that the statistical program is unable to calculate the regression coefficients due to singularity of the matrix.

¹⁵ The same conclusion can be made with respect to Imoisili (1989).

($X_2 \times X_3$ and $X_1 \times X_3$) are not.¹⁶ Second, this equation has a chance of being overspecified, which means that extra predictor variables are unnecessarily included. Overspecification of the model leads to an increased standard error of the regression coefficients (Cryer & Miller, 1991, p. 639), influencing the significance test of the coefficient. Harrison states that:

β_4 , as the (highest order) interaction term, is both stable and interpretable for its *standard error and significance test*. (Harrison, 1992, p. 11; emphasis added.)

Although the interaction term, standard error and significance test may be stable, they are hardly interpretable *if* the model is affected by an unnecessary overspecification. A standard error influenced by overspecification is not interpretable, no matter how stable it is. In this case, therefore, the significance of coefficient β_4 is not interpretable. Moreover, the partial derivative of this equation ($\partial Y/\partial X_2$) is:

$$\partial Y/\partial X_2 = \beta_2 + \beta_4 X_1 + \beta_5 X_3 \quad (9)$$

This reveals that the relationship between X_2 and Y is not a linear function of X_1 , but a linear function of both X_1 and X_3 . As a result, the hypothesis remains unanswered because the coefficient β_4 is not a measure of the hypothesized moderating effect of X_1 on the relationship between X_2 and Y .¹⁷

Apart from the use of multiple two-way interactions, MRA can be used to analyze n -way interactions. As an example of higher-order interactions, a three-way interaction generally has the format of Eq. (10):

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_1 \times X_2 + \beta_5 X_1 \times X_3 + \beta_6 X_2 \times X_3 + \beta_7 X_1 \times X_2 \times X_3 + \varepsilon \quad (10)$$

In conformity with the interpretation of the product term for two-way interactions, the product term ($X_1 \times X_2 \times X_3$) in Eq. (10) represents the three-way interaction among the (three) independent variables. The issues discussed above regarding use and interpretation of MRA for two-way interactions are also important for three-way interactions. Regarding the inclusion of main effects in MRA, for a three-way interaction equation, all two-way interactions should be included, and the user should observe that the coefficients obtained for these two-way interactions are not directly interpretable (Cohen & Cohen, 1983, p. 348).^{18,19} In general, for an n -way interaction, the main effects and all possible interactions of a lower-than- n order should be included. The difference between two-way and higher-order interactions

¹⁸ Once again, 'not interpretable' means that the coefficients found in the interaction model are not the same as those found in the main-effects-only model. Although the main effects are not interpretable in a three-way interaction equation, the two-way interactions are *if* the variables are centered. The interpretation is in conformity with the interpretation stated in previous sections, i.e. β_4 is the interaction effect $X_1 \times X_2$ on Y at the (sample) average of X_3 .

¹⁹ Brownell and Hirst (1986) use a three-way interaction equation and interpret a main effect [*Participation (P)*], a two-way interaction [*P* × *Budget Emphasis (B)*], and the three-way interaction [*P* × *B* × *Task Uncertainty (T)*]. The authors assume that every variable in the equation is interpretable. However, as was shown before, the coefficients of the main effects in a three-way interaction are normally not interpretable. Because only variable P is centered, only the two-way interaction without component P can be interpreted. This means that only the interaction of $B \times T$ can be interpreted. Brownell and Hirst (1986) on the other hand, interpret the coefficient of $P \times B$, which is statistically incorrect. Conclusion of the regression results should only have been based on the two-way interaction $B \times T$ and the three-way interaction ($P \times B \times T$). More problematic even is the analysis of Imoisili (1989, p. 330). Imoisili examines two-way interactions and uses a three-way interaction equation [see Eq. (10) above], where $Y = \text{a.o. Performance}$; $X_1 = \text{Budget style}$; $X_2 = \text{Task Interdependence}$; and $X_3 = \text{Task Uncertainty}$. The author is interested in the coefficients β_4 and β_5 . Their interpretation is hindered because of, among others, the use of raw scores instead of centred variables. The results therefore do not provide support for, nor allow rejection, of the hypothesis.

¹⁶ Harrison (1992) refers to both Schoonhoven (1981) and Jaccard et al. (1990) for using the multiple interaction equation. Schoonhoven's statistical analysis is incorrect, and referring to her analysis may have caused the incorporation of similar shortcomings. The reference to Jaccard et al. (1990, pp. 40–41) is intriguing, since here the importance of theory when using multiple interactions is stressed. Harrison seems to selectively use this passage to only eliminate the theoretically irrelevant three-way interaction (1992, p. 11), while leaving two theoretically irrelevant two-way interactions.

¹⁷ The same line of reasoning applies to Harrison (1993).

therefore does not lie in its mathematics or statistics, but in the interpretation of its meaning. Cohen and Cohen (1983, p. 306) note in this respect:

The fact that the mathematics can rigorously support the analysis of interactions of high order, however, does not mean that they should necessarily be constructed and used: interactions greater than three-way are most difficult to conceptualize, not likely to exist, and are costly in statistical inference (...)

However, even for a three-way interaction the problems noted by Cohen and Cohen (1983) exist. First, Schmidt and Hunter (1978) and Champoux and Peters (1987), for example, note that the sample sizes typical in (organizational) research lack the power to find the hypothesized interactions of a high-order. Budgetary studies are no exception, with sample sizes often far below 100 (cf. Young, 1996). Second, the difficulty of conceptualizing three-way interactions becomes clear if one considers that a three-way interaction means that a two-way interaction is a function of a third variable. In terms of Eq. (10), this means that a significant coefficient of the three-way interaction term (β_7) indicates that the interaction effect of X_1 and X_2 on Y is a function of X_3 . However, because of the symmetry mentioned earlier, a three-way interaction therefore simultaneously expresses:

- (a) the moderating effect of X_3 on the $X_1 \times X_2$ interaction affecting Y ;
- (b) the moderating effect of X_2 on the $X_1 \times X_3$ interaction affecting Y ; and,
- (c) the moderating effect of X_1 on the $X_2 \times X_3$ interaction affecting Y .

The complexity is evident if it is further considered that (a) implies that X_3 is ‘moderating the moderating effect’ that X_2 has on the relationship between X_1 and Y , and that also here the symmetry applies that was mentioned earlier. Because of the complexity, it seems even more important here that theory should dictate which variables are labeled the moderators.

A three-way interaction is graphically illustrated in Fig. 4, which depicts a three-way interaction for

one continuous and two dichotomous independent variables. Panel A and B both show a two-way interaction, and illustrate how the relationship between X_1 and Y is different for high and low values of X_2 . Panel A shows the two-way interaction for low values of X_3 . Panel B shows the two-way interaction for high values of X_3 . The three-way interaction signifies the difference between the two changes in slope for high and low values of X_3 . A significant three-way interaction does *not* indicate that a two-way interaction is significant for some values of X_3 , and not for others. Indeed, a three-way interaction can be significant, both when the ‘underlying’ two-way interactions are significant, and when they are not.²⁰ Note that here and in many (graphical) examples in this paper, the analysis is simplified by using *dichotomous* variables. For three continuous variables, the graphical depiction and explanation of three-way interactions are far more complex. In sum, Cohen and Cohen (1983, pp. 347–348) therefore advocate a restricted use of higher-order interactions, and state:

No interaction set should be included in the IV’s (*independent variables*) unless it is seriously entertained on substantive grounds (...). This requires as a minimum condition that it be understood by the investigator and on practical grounds that it can be clearly explicable to his audience.

Of the 28 papers analyzed, six papers specifically study three-way interactions (Brownell & Dunk, 1991; Brownell & Hirst, 1986; Dunk, 1993; Gul & Chia, 1994; Harrison, 1993; Lau et al., 1995). Three of these six papers decompose the three-way interaction into two-way interactions (Brownell & Dunk, 1991, pp. 700–701; Brownell & Hirst, 1986, pp. 247–248; Lau et al., 1995, p. 372). In this way the authors can ascertain that the significant three-way interaction is indeed a ‘real’

²⁰ Thus, a *significant universal* relationship may exist between the independent and dependent variables, while the three-way interaction is also significant. An example of (hypothetical) data that exhibit this relationship is available from the authors upon request.

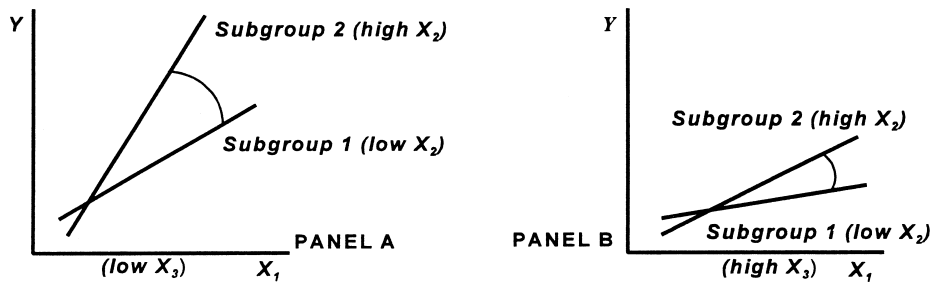


Fig. 4. Three-way interactions for two dichotomous moderators.

relationship and not fallacious.²¹ As stated before, a significant three-way interaction does not contain any information about the significance of the underlying two-way interactions. To establish the latter, the additional analysis is required.²² The other three papers mentioned (Dunk, 1993; Gul & Chia, 1994; Harrison, 1993) do not check the significance of the two-way interaction effects underlying the three-way interactions found. Solely based on the significance of the three-way interaction, the authors draw conclusions with respect to the underlying relationship at the level of two-way interactions (Dunk, 1993, p. 407; Harrison, 1993, p. 333) and even at the level of main effects (Gul & Chia, 1994, p. 422), Gul & Chia (1994), for example, measure the three-way interaction of *Management Accounting System Scope*, *Perceived Environmental Uncertainty* and *Decentralization on Managerial Performance*. They use the partial derivative of the three-way interaction Eq. (10) to analyze two-way interactions and main effects. The result of their analysis is theoretically uninterpretable, since they do not establish whether two-way interactions exist at all. The same conclusion holds for the papers of Dunk (1993) and Harrison (1993).

²¹ Lau et al. (1995) provide a clear example of such a decomposition. A flaw in their analysis, however, is that they define a null hypothesis and alternative hypotheses (pp. 363–364), which are not mutually exclusive. This means that the statistical results *could* have supported both the null hypothesis and the alternative hypotheses.

²² Brownell and Dunk (1991) do the additional analysis to confirm the expected form and sign of the two-way interactions (p. 701). Finding a significant three-way interaction does not warrant such specific expectations.

It seems clear that Cohen and Cohen's (1983) warning concerning higher-order interactions is not heard in the budgetary research paradigm. A reviewer of Gul and Chia's paper suggested the use of a *four-way* interaction (1994, footnote 10). It is easy to consider the problems and complexities associated with this format.

4.5. Interaction, effect size, and the Johnson–Neyman technique

A further issue related to the interpretation of the outcomes of MRA is that a (significant) coefficient of the interaction only contains information about changes in the relationship between variables, and does not contain information about the optimal value of the dependent variable (i.e. effect size, see Champoux & Peters, 1987). This is illustrated in Fig. 5. In panel A, Y has the highest (lowest) value when both X1 and X2 are high (low). In contrast, in panel B, Y has the highest (lowest) value when X1 is high (low) and X2 is low (high). Note that in both cases the interactions ($X_1 \times X_2$) are equal with respect to both direction and size. This means that in both cases, an increase in the value of X2 has an equal positive effect on the form of the relationship between X1 and Y. The difference between the cases is due to a main effect of the moderating variable on the dependent variable (cf. Kren & Kerr, 1993). The proper interpretation of a positive interaction therefore is *not* that Y achieves the highest values for the highest values of X1 and X2, but that for higher values of X1, X2 has a more positive effect on Y. Hence, MRA *cannot* be used to test expectations about

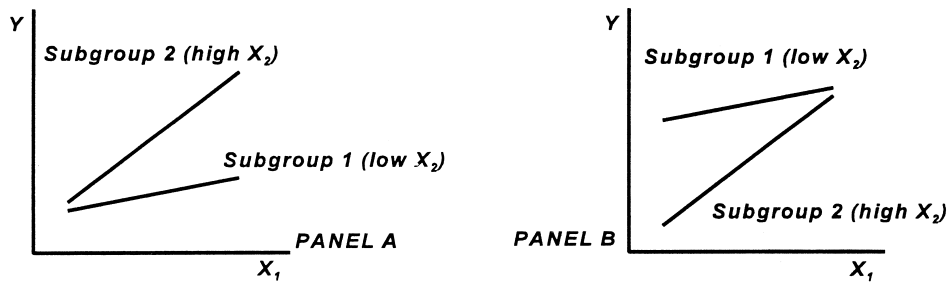


Fig. 5. Interaction contains no information on value independent variable.

the values of X_1 and X_2 for which Y will have the highest value. This is the consequence of MRA testing the significance of the interaction effect, and not testing for a *combined* effect of the main effects and the interaction effect on the dependent variable.

The difference between a significant interaction and effect size is not always recognized in the papers examined. For example, Dunk (1992) hypothesizes that:

(...) the higher (lower) the level of manufacturing process automation *and* the higher (lower) the reliance on budgetary control, the higher will be production subunit performance (p. 198; emphasis added).

This hypothesis states that high/high and low/low combinations of the independent variables maximize the dependent variable. Dunk (1992) uses MRA to test this hypothesis and thus incorrectly assumes that the significant interaction contains information about effect size. Further, Brownell (1983), Brownell and Hirst (1986), Mia (1989), Dunk (1989, 1990, 1993) and Brownell and Dunk (1991) also assume that a significant interaction means that a certain combination of variables maximizes the dependent variable. For example, Dunk (1993) states the following about the significance of the coefficient of the three-way interaction term (b_7):

As b_7 (...) is significant and negative, it appears that slack is low when participation,

information asymmetry, and budget emphasis are *all* high (pp. 405–406; emphasis added).

Such a statement is not only incorrect, the question of the ‘highest Y ’ is likely to be *irrelevant* in testing contingency models, especially when the main effects are due to uncontrollable, exogenous contingency factors.

Two papers apply a formal analysis of effect sizes by using the so-called Johnson–Neyman technique (Brownell, 1982a; Lau et al., 1995). This technique can be used to find the ‘region of significance’ for the difference between the effects of different values of the moderator at a given value of the independent variable (Pedhazur & Pedhazur-Schmelkin, 1991). In other words, the technique provides a measure to establish whether the effect of the moderator is ‘large enough’ to lead to significant different values of the dependent variable. The Johnson–Neyman technique can be illustrated more clearly by returning to Fig. 5. The question is whether for a given value of X_1 , there is a significant difference between the value of Y for subgroup 1 and 2 (i.e. for $X_2 = \text{low}$ and $X_2 = \text{high}$). In panel A, the ‘region of significance’ will relate to higher values of X_1 , since the difference between the two regression lines increases with increases in X_1 . In panel B, on the other hand, the ‘region of significance’ will relate to lower values of X_1 . Despite this difference in ‘region of significance’, the interactions are equal with respect to both direction and size, as stated before. The use of the Johnson–Neyman technique in MRA is therefore questionable, since it plays no

role in the test of hypotheses of the interaction format. Since the interaction contains no information on effect sizes (i.e. values of the dependent variable), the results of this technique are of no relevance to the interpretation of the interaction effect. The formula used to find the 'region of significance' is not only determined by the moderating effect (i.e. effect on slope) of the contingency variable, but also by its effect on the *intercept*. The region of significance could be very large even if the interaction effect is very small and insignificant, because of a large difference in the intercept (and vice versa). Further, looking at Eqs. (2a) and (2b), one sees that the difference between intercepts is influenced by the 'main effect of the moderator' (cf. Kren & Kerr, 1993). As a result, the Johnson–Neyman technique mixes main effects and moderating effects, and thus seems of little value to explore the nature of the interaction effect alone, as stated and done by Brownell (1982a) and Lau et al. (1995).

4.6. Interaction and (non-)monotonicity

Both in formulating and in testing contingency hypotheses of the interaction format, it is important to consider the (non-)monotonicity of the hypothesized relationship. A statistically significant coefficient of the interaction term does not contain information about whether the relationship found is monotonic or non-monotonic, nor does contingency theory have an a priori 'preference' for monotonic or non-monotonic relationships. Since, however, the substantive implications are different for monotonic and non-monotonic relationships, studies should explicitly state whether the aim is to investigate and test (non-)monotonicity.^{23,24} Six of the 28 papers use

the method of the partial derivative to analyze the (non-)monotonicity of relationships found (Govindarajan & Gupta, 1985; Gul & Chia, 1994; Harrison, 1993; Lau et al., 1995²⁵; Mia, 1988, 1989). For example, Govindarajan and Gupta (1985) measure the partial derivative of two two-way interactions and provide graphs of these equations. The resulting two graphs are examples of an almost perfect non-monotonic relationship in which the line crosses the X-axis near zero. The conclusion is that for one extreme value of the moderator (i.e. *Strategy*) the organization will be effective if, here, accounting information is used in performance evaluation, while being ineffective for the other extreme value of the moderator. Harrison (1993) and Gul and Chia (1994) both measure the partial derivative of the three-way interaction Eq. (10). However, as the above analysis shows that their results are uninterpretable because of a misspecified model, the conclusions about non-monotonicity are invalid as well.

The graph of the partial derivative is one test for non-monotonic effects but non-monotonicity can also be measured by a regression per subgroup. Four of the 28 papers use subgroup regression analysis and all indicate that a non-monotonic relationship exists (Brownell, 1983, 1985; Brownell & Merchant, 1990; Mia & Chenhall, 1994). A further strong point of these studies is that, except for Brownell and Merchant (1990), they all measure the statistical significance for the different subgroups, which allows a better understanding of the higher-order interaction. In the majority of papers, however, the issue of (non-)monotonicity is not addressed.

5. Summary of findings, conclusions and implications

The evidence in the previous sections leads to the initial conclusion that the use of MRA in the papers reviewed is seriously flawed, caused by the uncritical application of this statistical technique and too little knowledge of its specific requirements and underlying assumptions. Table 2 presents an overview of the findings. Generally, it appears that six major types of errors in MRA use

²³ Only Mia (1988, 1989) explicitly states a non-monotonic form hypothesis.

²⁴ Recall that an important reason for budgetary research to adopt a contingency perspective were the opposite results of Hopwood (1972) and Otley (1978). Hopwood (1972) found a *positive* effect of budget emphasis, whereas Otley (1978) found a *negative* effect. A contingency hypothesis that attempts to explain these *opposite* effects of budget emphasis can only do so by predicting a *non-monotonic* interaction effect.

²⁵ Lau et al. (1995) only mention the result of the non-monotonic test in a footnote (p. 374).

Table 2
Summary and overview of findings

Major types of errors in MRA use	Number of articles
1. Format of statistical test not in conformity with hypothesis	21 (75%)
2. Faulty use of tests for interactions of the 'strength' type	4 (14%)
3. Incorrect interpretation of main effects	8 (29%)
4. Incorrect specification of the MRA equation	4 (14%)
5. Incorrect use of higher-order interactions	3 (11%)
6. Incorrect conclusions about effect sizes from MRA	10 (36%)
Number of articles containing at least one of the above errors	27 (96%)

frequently occur in budgetary studies. These are: (1) format of statistical test not in conformity with hypothesis; (2) faulty use of tests for interactions of the 'strength' type when hypothesizing interactions of the 'form' type; (3) incorrect interpretation of main effects; (4) incorrect specification of the MRA equation; (5) incorrect use of higher-order interactions equations to test lower-order interactions; and, (6) incorrect conclusions about effect sizes from MRA. Overall, 27 of the 28 papers (96%) exhibit at least one of the above errors.²⁶

Although the summary of findings reveals that only one paper appears free from errors, this does not imply that the statistical results presented in all other papers are meaningless, nor that the conclusions drawn and presented in these papers are not supported by the data. Regarding the former, although researchers may have applied MRA incorrectly, and may have interpreted the MRA results incorrectly, it may be that the statistical results presented in these studies are interpretable and useful. Therefore, an additional analysis was done to evaluate the interpretability of the statistical results *as presented*. This further analysis reveals that the statistical results of 12 of the 28 papers (43%) can still not be interpreted. For these papers, the uninterpretability of statistical results is due to: (a) the use of subgroup correlation analysis instead of the required MRA (Govindarajan, 1984; Merchant, 1981, 1984, 1990); (b) the incorrect specification of the MRA equation (Brownell, 1982b; Frucot & Shearon,

1991; Harrison, 1992, 1993; Hirst, 1983; Imoisili, 1989); and (c) the deficient analysis of a three-way interaction (Dunk, 1993; Gul & Chia, 1994).

Regarding the latter, the analysis in the present paper does also not prove that the conclusions drawn and presented in the reviewed papers are incorrect and unsupported by the data. It may be that the statistical results presented in these papers are robust and insensitive to the analytical flaws and model misspecifications found. To check the robustness of the results, another additional analysis would be required. Such an analysis would imply the re-analysis of the original data using MRA in conformity with the methodology and a subsequent comparison of the results with those presented in the original papers. Thus far, the authors have not been able to conduct this test.²⁷

²⁷ To conduct such an analysis, the authors asked, at the outset of the paper, for the data from two recent papers in the sample. These two papers explicitly stated that the data 'were available upon request'. The authors of the two papers were approached by both regular mail and e-mail. The letter and e-mail stated the subject of the present paper and the purpose of the request, which was to analyze the data for strictly methodological reasons. The results of this request were disappointing. The author(s) of one paper replied that the data were lost due to a move to a new university. The author(s) of the second paper did not reply at all. After these two 'answers', data were asked from a third budgetary control paper. This was not included in the sample (since it did not test 'interaction'), but was comparable and deemed useful for the additional data-analysis. Also here it said that the data were available. In this case the author quickly replied but stated that the data were lost due to a 'computer crash'. Overall, this raises suspicion about the actual data availability and, consequently, of the value of a 'data availability policy'. Although the aim of this paper is not to investigate the effectiveness of data availability policy proposed by some journals, such an investigation does seem in order.

²⁶ The paper of Govindarajan and Gupta (1985) does not contain any errors with respect to the application and interpretation of MRA.

In sum, the findings in this paper provide clear evidence that the use of statistics in the budgetary contingency literature does not indicate a high level of technical quality. Many studies show too little knowledge of the characteristics and pitfalls of MRA, and do not display the expertise and care required in the interpretation of its outcomes. Moreover, budgetary studies contain little rigor in their use of contingency theory, since it is also found that many studies do not provide a good link between the verbal (substantive) format of the hypothesis and the statistical format subsequently used to test the hypothesis. The reason for this apparent negligence is not obvious, but it may be an additional ground to be critical at *theoretical* advancement in this area of the literature (cf. Chapman, 1997; Lindsay & Ehrenberg, 1993; Hartmann, in press; Young, 1996). Earlier Briers and Hirst (1990, p. 385) have sharply criticized the underdevelopment of contingency theory in many budgetary and RAPM studies. They stated:

Of particular concern is the inclusion of variables in hypothesis with little supporting explanation. For example, some studies use box diagrams with arrows indicating causally related variables. Although this is a parsimonious way of communicating connections, the supporting argument in some studies is only suggestive (...).

This apparent lack of ambition to develop a true contingency theory of management accounting was noted before by Otley, who suggested that in many studies:

... [t]he contingency approach is invoked, so it seems, in order to cover up some of the embarrassing ambiguities that exist in the universalistic approach (Otley, 1980, p. 414).

Indeed, many of the papers that have appeared since then still suffer from the defects at which Otley was hinting. These two main conclusions provide ample reason to be worried about the current state of budgetary contingency research for at least three reasons. First, the analysis only

included papers from high-quality accounting journals. Second, the analysis referred to an area of research considered to be of great importance to the broad area of management accounting research. Third, the analysis examined a research methodology which has become typical for this and related research fields. The dangers of the impact and persistence of errors in MRA found for the state of knowledge in this specific field of research are large given the lack of successful replication studies here (cf. Lindsay & Ehrenberg, 1993), and the lack of large sample studies (Lindsay, 1995).

The main implication for future research is that major advancements in this field can be made. Regarding the technical failures in MRA, the findings in this study provide strong support for earlier pleas for the improvement of the methodological quality of management accounting research (cf. Lindsay & Ehrenberg, 1993; Lindsay, 1995; Young, 1996). In itself, MRA is a method that is well-regarded and well-described in the literature. Regarding the flaws that affect both MRA and contingency theory, authors should strive for better and more explicit articulations of contingency hypotheses. Moreover, additional care is required in linking the form of the theoretical proposition with the format of the statistical test. This could also mean an increased focus on other than simply ‘interaction’ types of contingency fit (see e.g. Venkatraman, 1989). Such more consciously matched theories, hypotheses and tests are the necessary ingredients to develop a ‘true’ contingency theory of management accounting (cf. Chapman, 1997). Since it is the theory that dictates the format of ‘contingency fit’, it should also be theory that dictates the appropriate way of testing ‘contingency fit’.

Appendix A

Selected forms and types of contingency fit

Appendix B

Overview of reviewed articles

Table A.1
 Examples of different types of contingency fit. The first column states the 'type of fit'. The second column then presents a typical hypothesis which is formulated in accordance with the type of fit. Since many researchers in budgetary research are interested how the relationship between *RAPM* and *Performance (P)* is contingent on *Environmental Uncertainty (EU)*, the typical hypothesis is stated as an example using these three variables.

Type of fit	Typical hypothesis	Appropriate statistical test (criterion)	Selected references
Interaction (form, monotonic)	For lower <i>EU</i> , the effect of <i>RAPM</i> on <i>P</i> is more positive (more negative)	Moderated Regression Analysis (significant coefficient of the interaction term)	Southwood (1978) Arnold (1982, 1984) Jaccard et al. (1990)
Interaction (form, non-monotonic)	For higher values of <i>EU</i> , <i>RAPM</i> will positively (negatively) affect <i>P</i> , for lower values of <i>EU</i> , <i>RAPM</i> will negatively (positively) affect <i>P</i>	Moderated Regression Analysis (significant coefficient of the interaction term plus partial derivative equals zero within range)	Schoonhoven (1981)
Interaction (strength)	The relationship between <i>RAPM</i> and <i>P</i> is stronger (better predicted) when <i>EU</i> is low than when <i>EU</i> is high	Subgroup correlation analysis (significant difference in correlation coefficients)	Arnold (1982, 1984)
Mediation	The relationship between <i>EU</i> and <i>P</i> is explained by an indirect effect whereby <i>EU</i> reduces <i>RAPM</i> , which in turn increases <i>P</i>	Path analysis (significant path coefficients)	Venkatraman (1989)
Selection	When <i>EU</i> is higher, <i>RAPM</i> is higher	Correlation analysis (significant correlation coefficient)	Drazin and Van de Ven (1985)
Matching	Given the value of <i>EU</i> , there is a unique value of <i>RAPM</i> that maximizes <i>P</i> , deviations from this relationship in either direction reduces the value of <i>P</i>	Multiple regression analysis including a 'matching' term, e.g. $Y = X + Z + X - Z $ (significance of coefficient of the matching variable)	Johns (1981) Bedeian et al. (1994) Venkatraman (1989)
Systems approach	For each contextual situation there is an ideal control system, namely (<i>RAPM</i> , <i>EU</i> , other budgetary and/or contingency variables); deviations from this ideal profile reduce <i>P</i>	Correlation between Euclidian distance and effectiveness (significant correlation coefficient)	Drazin and Van de Ven (1985)

Table A.2
Overview of reviewed articles

Author	Research model		Contingency fit		Dependent variables	Hypothesized	Tested
	Contingency variables	Budgetary variables	Budgetary variables	Contingency fit			
Merchant (1981) TAR ^a	Size Diversity Degree of decentralization	Budgeting system	Management —motivation —attitude Organizational performance	Selection Interaction (form)	Management —motivation —attitude Organizational performance	Selection Interaction (form)	Selection Interaction (strength)
Brownell (1982a) JAR ^b		Supervisory evaluation style Budgetary participation	Performance Job satisfaction	Null hypothesis (interaction)	Performance Job satisfaction	Null hypothesis (interaction)	Interaction (form)
Brownell (1982b) TAR	Locus of control	Budgetary participation	Performance Job satisfaction	Null hypothesis (interaction)	Performance Job satisfaction	Null hypothesis (interaction)	Matching
Brownell (1983) JAR		Management by exception Budgetary participation	Motivation	Null hypothesis (interaction)	Motivation	Null hypothesis (interaction)	Interaction (form)
Hirst (1983) JAR	Task uncertainty	RAPM	Dysfunctional behavior	Curvilinear	Dysfunctional behavior	Curvilinear	Curvilinear Interaction (strength) Interaction (form)
Govindarajan (1984) AOS ^c	Environmental uncertainty	Performance evaluation and reward system	Effectiveness	Selection Interaction (strength)	Effectiveness	Selection Interaction (strength)	Selection Interaction (strength)
Merchant (1984) AOS	Production technology Market factors Organizational characteristics	Budgeting system	Performance	Selection Interaction (form)	Performance	Selection Interaction (form)	Selection Interaction (strength)
Brownell (1985) JAR	Functional area	Budgetary participation RAPM	Managerial performance	Null hypothesis (interaction)	Managerial performance	Null hypothesis (interaction)	Interaction (form)
Govindarajan and Gupta (1985) AOS	Business unit strategy	Bonus criteria Determination of bonus	Effectiveness	Interaction (form)	Effectiveness	Interaction (form)	Interaction (form), plus test for (non-) monotonic effects
Brownell and Hirst (1986) JAR	Task uncertainty	Budget emphasis Budgetary participation	Performance Job related tension	None	Performance Job related tension	None	Interaction (form)

continued overleaf

Table A.2—cont'd

Author	Research model		Contingency fit		
	Contingency variables	Budgetary variables	Dependent variables	Hypothesized	Tested
Chenhall (1986) TAR	Authoritarianism	Budgetary participation	Job satisfaction with budgets	Null hypothesis (interaction)	Interaction (form)
Mia (1988) AOS	Attitude	Budgetary participation	Performance Motivation	Interaction (non-monotonic form)	Interaction (non-monotonic form)
Dunk (1989) AOS		Budget emphasis Budgetary participation	Performance	Null hypothesis (interaction)	Interaction (form)
Mia (1989) AOS	Job difficulty	Budgetary participation	Performance Motivation	Interaction (non-monotonic form)	Interaction (non-monotonic form)
Imoisili (1989) AOS	Task interdependency Task uncertainty	Budget style	Job stress Performance Attitudes towards budgets	Interaction (form)	Interaction (form)
Brownell and Merchant (1990) JAR	Product standardization Manufacturing process auto-mation	Budgetary participation Flexible budgeting	Departmental performance	None	Interaction (form)
Dunk (1990) AOS	Agreement on evaluation criteria	Budgetary participation	Managerial performance	Null hypothesis (interaction)	Interaction (form)
Hirst and Lowy (1990) AOS		Budgetary goal difficulty Budgetary feedback	Budgetary performance Overall job performance	Interaction	Interaction (form)
Merchant (1990) AOS	Environmental uncertainty Supervisory consideration Profit center strategy	Pressure to meet financial	Dysfunctional behavior targets	Interaction	Interaction (strength)

Brownell and Dunk (1991) AOS	Task uncertainty (task difficulty and task variability)	Budgetary participation Budget emphasis	Managerial performance	None	Interaction (form)
Frurot and Shearon (1991) TAR	Locus of control	Budgetary participation	Managerial performance Job satisfaction	Null hypothesis (interaction)	Matching
Dunk (1992) AOS	Process automation	Budget emphasis	Production subunit performance	High/high and low/low combinations of the independent variables lead to high performance	Interaction (form)
Harrison (1992) AOS	Culture	Participation Budget emphasis	Job related tension Job satisfaction	Null hypothesis (interaction)	Interaction (form)
Dunk (1993) TAR	Information asymmetry	Budgetary participation Budget emphasis	Budgetary slack	Null hypothesis (interaction)	Interaction (form)
Harrison (1993) AOS	National culture Personality	<i>RAPM</i>	Job related tension Job satisfaction	Null hypothesis (interaction)	Interaction (form), plus test for (non-) monotonic effects
Gul and Chia (1994) AOS	Perceived environmental uncertainty	MAS Decentralization	Managerial performance	Interaction (form)	Interaction (form)
Mia and Chenhall (1994) AOS	Functional differentiation	Use of broad scope MAS	Managerial performance	Null hypothesis (interaction)	Interaction (form)
Lau et al. (1995) AOS	Task uncertainty	Budget emphasis Budgetary participation	Job related tension Managerial performance	Interaction (form)	Interaction (form)

^aTAR, *The Accounting Review*.

^bJAR, *Journal of Accounting Research*.

^cAOS, *Accounting, Organizations and Society*.

References

- Argyris, C. (1952). *The impact of budgets on people*. Ithaca, New York: The Controllership Foundation.
- Arnold, H. J. (1982). Moderator variables: A clarification of conceptual, analytic, and psychometric issues. *Organizational Behavior and Human Performance*, 29, 143–174.
- Arnold, H. J. (1984). Testing moderator variables hypotheses: A reply to Stone and Hollenbeck. *Organizational Behavior and Human Performance*, 34, 214–224.
- Arnold, H. J., & Evans, M. G. (1979). Testing multiplicative models does not require ratio scales. *Organizational Behavior and Human Performance*, 24, 41–59.
- Bedeian, A. G., Day, D. V., Edwards, J. R., Tisak, J., & Smith, C. S. (1994). Difference scores: rationale, formulation, and interpretation. *Journal of Management*, 20, 673–698.
- Briers, M., & Hirst, M. K. (1990). The role of budgetary information in performance evaluation. *Accounting, Organizations and Society*, 15, 373–398.
- Brown, L. D., & Huefner, R. J. (1994). The familiarity with and perceived quality of accounting journals: Views of senior accounting faculty in leading U.S. MBA programs. *Contemporary Accounting Research*, 11, 223–250.
- Brownell, P. (1982a). The role of accounting data in performance evaluation, budgetary participation and organizational effectiveness. *Journal of Accounting Research*, 20, 12–27.
- Brownell, P. (1982b). A field study examination of budgetary participation and locus of control. *The Accounting Review*, 57, 766–777.
- Brownell, P. (1983). The motivational impact of management-by-exception in a budgetary context. *Journal of Accounting Research*, 21, 456–472.
- Brownell, P. (1985). Budgetary systems and the control of functionally differentiated organizational activities. *Journal of Accounting Research*, 23, 502–512.
- Brownell, P. (1995). *Research methods in management accounting*. *Accounting Research Methodology*, Monograph No. 2, Melbourne: Coopers and Lybrand.
- Brownell, P., & Dunk, A. S. (1991). Task uncertainty and its interactions with budgetary participation and budget emphasis: Some methodological issues and empirical investigation. *Accounting, Organizations and Society*, 16, 693–703.
- Brownell, P., & Hirst, M. K. (1986). Reliance on accounting information, budgetary participation, and task uncertainty: Tests of a three-way interaction. *Journal of Accounting Research*, 24, 241–249.
- Brownell, P., & Merchant, K. A. (1990). The budgetary and performance influences of product standardization and manufacturing process automation. *Journal of Accounting Research*, 28, 388–397.
- Champoux, J. E., & Peters, W. S. (1987). Form, effect size and power in moderated regression analysis. *Journal of Occupational Psychology*, 60, 243–255.
- Chapman, C. S. (1997). Reflections on a contingent view of accounting. *Accounting, Organizations and Society*, 22, 189–205.
- Chenhall, R. H. (1986). Authoritarianism and participative budgeting: A dyadic analysis. *The Accounting Review*, 61, 263–272.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale: Erlbaum.
- Cronbach, L. J. (1987). Statistical tests for moderator variables: Flaws in analyses recently proposed. *Psychological Bulletin*, 102, 414–417.
- Cryer, J. D., & Miller, R. B. (1991). *Statistics for business: Data analysis and modelling*. Boston: PWS-Kent.
- Drazin, R., & Van de Ven, A. H. (1985). Alternative forms of fit in contingency theory. *Administrative Science Quarterly*, 30, 514–539.
- Dunk, A. S. (1989). Budget emphasis, budgetary participation and managerial performance. A note. *Accounting, Organizations and Society*, 14, 321–324.
- Dunk, A. S. (1990). Budgetary participation, agreement on evaluation criteria and managerial performance: A research note. *Accounting, Organizations and Society*, 15, 171–178.
- Dunk, A. S. (1992). Reliance on budgetary control, manufacturing process automation and production subunit performance: A research note. *Accounting, Organizations and Society*, 17, 195–203.
- Dunk, A. S. (1993). The effect of budget emphasis and information asymmetry on the relation between budgetary participation and slack. *The Accounting Review*, 68, 400–410.
- Dunlap, W. P., & Kemery, E. R. (1987). Failure to detect moderating effects: is multicollinearity the problem? *Psychological Bulletin*, 102, 418–420.
- Fisher, J. (1995). Contingency-based research on management control systems: Categorization by levels of complexity. *Journal of Accounting Literature*, 14, 24–53.
- Frucot, V., & Shearon, W. T. (1991). Budgetary participation, locus of control, and Mexican managerial performance and job satisfaction. *The Accounting Review*, 66, 80–99.
- Govindarajan, V. (1984). Appropriateness of accounting data in performance evaluation: An empirical investigation of environmental uncertainty as an intervening variable. *Accounting, Organizations and Society*, 9, 125–135.
- Govindarajan, V., & Gupta, A. K. (1985). Linking control systems to business unit strategy: Impact on performance. *Accounting, Organizations and Society*, 10, 51–66.
- Gul, F. A., & Chia, Y. M. (1994). The effects of management accounting systems, perceived environmental uncertainty and decentralization on managerial performance: A test of a three-way interaction. *Accounting, Organizations and Society*, 19, 413–426.
- Harrison, G. L. (1992). The cross-cultural generalizability of the relation between participation, budget emphasis and job-related attitudes. *Accounting, Organizations and Society*, 17, 1–15.
- Harrison, G. L. (1993). Reliance on accounting performance measures in supervisor evaluative style: The influence of national culture and personality. *Accounting, Organizations and Society*, 18, 319–339.
- Hartmann, F. G. H. (1999). The appropriateness of RAPM: toward the further development of theory. *Accounting, Organizations and Society*, in press.

- Hirst, M. K. (1983). Reliance on accounting performance measures, task uncertainty and dysfunctional behavior: Some extensions. *Journal of Accounting Research*, 21, 596–605.
- Hirst, M. K., & Lowy, S. M. (1990). The linear additive and interactive effects of budgetary goal difficulty and feedback on performance. *Accounting, Organizations and Society*, 15, 425–436.
- Hopwood, A. G. (1972). An empirical study of the role of accounting data in performance evaluation. *Journal of Accounting Research*, 10, 156–182.
- Imoisili, O. A. (1989). The role of budget data in the evaluation of managerial performance. *Accounting, Organizations and Society*, 14, 325–335.
- Jaccard, J., Turrisi, R., & Wan, C. K. (1990). *Interaction Effects in Multiple Regression*. Newbury Park: Sage.
- Johns, G. (1981). Difference score measures of organizational behavior variables: a critique. *Organizational Behavior and Human Performance*, 27, 443–463.
- Kren, L., & Kerr, J. L. (1993). The effect of behavior monitoring and uncertainty on the use of performance-contingent compensation. *Accounting and Business Research*, 23, 159–168.
- Kren, L., & Liao, W. M. (1988). The role of accounting information in the control of organizations: A review of the evidence. *Journal of Accounting Literature*, 7, 280–309.
- Lau, C. M., Low, L. C., & Eggleton, I. R. C. (1995). The impact of reliance on accounting performance measures on job-related tension and managerial performance: Additional evidence. *Accounting, Organizations and Society*, 20, 359–381.
- Lindsay, R. M. (1995). Reconsidering the status of tests of significance: An alternative criterion of adequacy. *Accounting, Organizations and Society*, 20, 35–53.
- Lindsay, R. M., & Ehrenberg, A. S. C. (1993). The design of replicated studies. *The American Statistician*, 47, 217–228.
- Mia, L. (1988). Managerial attitude, motivation and the effectiveness of budget participation. *Accounting, Organizations and Society*, 13, 465–475.
- Mia, L. (1989). The impact of participation in budgeting and job difficulty on managerial performance and work motivation: A research note. *Accounting, Organizations and Society*, 14, 347–357.
- Mia, L., & Chenhall, R. H. (1994). The usefulness of management accounting systems, functional differentiation and management effectiveness. *Accounting, Organizations and Society*, 19, 1–13.
- Merchant, K. A. (1981). The design of the corporate budgeting system: Influences on managerial behavior and performance. *The Accounting Review*, 56, 813–829.
- Merchant, K. A. (1984). Influences on departmental budgeting: An empirical examination of a contingency model. *Accounting, Organizations and Society*, 9, 291–307.
- Merchant, K. A. (1990). The effects of financial controls on data manipulation and management myopia. *Accounting, Organizations and Society*, 15, 297–313.
- Otley, D. T. (1978). Budget use and managerial performance. *Journal of Accounting Research*, 16, 122–149.
- Otley, D. T. (1980). The contingency theory of management accounting: Achievement and prognosis. *Accounting, Organizations and Society*, 5, 413–428.
- Pedhazur, E. J., & Pedhazur-Schmelkin, L. (1991). *Measurement, design and analysis: An integrated approach*. Hillsdale: Erlbaum.
- Schmidt, F. L., & Hunter, J. E. (1978). Moderator research and the law of small numbers. *Personnel Psychology*, 31, 215–232.
- Schoonhoven, C. B. (1981). Problems with contingency theory: Testing assumptions hidden within the language of contingency 'Theory'. *Administrative Science Quarterly*, 26, 349–377.
- Sharma, S., Durand, R. M., & Gur-Arie, O. (1981). Identification and analysis of moderator variables. *Journal of Marketing*, 18, 291–300.
- Shields, J. F., & Shields, M. D. (1998). Antecedents of participative budgeting. *Accounting, Organizations and Society*, 23, 49–76.
- Southwood, K. E. (1978). Substantive theory and statistical interaction: Five models. *American Journal of Sociology*, 83, 1154–1203.
- Stone, E. F., & Hollenbeck, J. R. (1984). Some issues associated with the use of moderated regression. *Organizational Behavior and Human Performance*, 34, 195–213.
- Venkatraman, N. (1989). The concept of fit in strategy research: Toward verbal and statistical correspondence. *Academy of Management Review*, 14, 423–444.
- Young, S. M. (1996). Survey research in management accounting: A critical assessment. In A. J. Richardson, (Ed.), *Research methods in accounting*. Vancouver: CGA Research Foundation.