



ELSEVIER

Mathematical Social Sciences 43 (2002) 485–504

mathematical
social
sciences

www.elsevier.com/locate/econbase

Enumerating and testing conjoint measurement models

George Karabatsos^{a,*}, James R. Ullich^b

^a*Louisiana State University Health Sciences Center, 1600 Canal Street, Suite 800, 70112 New Orleans, LA, USA*

^b*Southern Utah University, Utah, USA*

Received 1 May 2001; received in revised form 1 April 2002; accepted 1 April 2002

Abstract

A Monte Carlo procedure was used to generate three types of $3 \times 3 \times 2$ conjoint measurement structures, each type having an additive (respectively, distributive, dual-distributive) simple polynomial representation. The Monte Carlo results illustrate the restrictiveness of the axioms of joint independence (respectively, distributive cancellation, dual-distributive cancellation), and show how close each comes to characterizing the relevant simple polynomial representation. An empirical example illustrates how the generated measurement structures aid in the selection and testing of axiomatic models using order-restricted statistical inference.

© 2002 Elsevier Science B.V. All rights reserved.

Keywords: Conjoint measurement; Axiomatic measurement theory; Enumerating measurement structures; Axiom testing

1. Review of polynomial conjoint measurement

Polynomial conjoint measurement (PCM) is useful for measuring how a dependent variable changes as a function of two or more independent variables. Hence, PCM is similar in spirit to classical analysis-of-variance, although PCM offers a richer and more insightful description of the data structure (e.g. Krantz et al., 1971, hereafter called KLST; see also Krantz and Tversky, 1971). PCM identifies measurement axioms that allow for a detailed test of the hypotheses that a particular composition rule (a function of the independent variables) describes a given data structure (see KLST, 1971, Fig. 1, p. 345). Since these axioms specify qualitative inequality restrictions, PCM is particularly useful for the social sciences. This is because psychological variables (e.g. judgement,

*Corresponding author. Tel.: +1-504-568-4406; fax: +1-504-568-6905.

E-mail address: gkarab@lsuhsc.edu (G. Karabatsos).

attitudes, and ability) often can not be directly observed, and therefore data arising from these variables often can only be measured up to an ordinal scale.

We now summarize enough of KLST (KLST, 1971, Section 3.3) to set the PCM framework for our own results. We first define *simple polynomials*, and then describe the four possible classes of simple polynomials in three variables, namely, the additive, distributive, dual-distributive, and multiplicative composition rules.

Simple polynomials (Krantz and Tversky, 1971, p. 328) are characterized by the following: (a) any single variable is a simple polynomial; (b) if, and only if, two simple polynomials have no variables in common, then their sum and their product is also a simple polynomial. There are four classes of simple polynomials in three variables, yielding a total of eight such polynomials. To state those four, let X_1 , X_2 , and X_3 be three, real-valued, independent variables. The four composition rules (KLST, 1971, p. 328) are:

1. *Additive*: $X_1 + X_2 + X_3$.
2. *Distributive*: $(X_1 + X_2)X_3$, and two others by interchanging X_1 and X_3 or X_2 and X_3 .
3. *Dual-Distributive*: $X_1X_2 + X_3$, and two others by interchanging variables, as above.
4. *Multiplicative*: $X_1X_2X_3$.

Note that, for convenience, this study focuses on a special case¹ of PCM, where the scales of X_1 , X_2 , and X_3 are restricted to the positive real domain, Re^+ . In this case, the additive and the multiplicative composition rules are (ordinally) equivalent (e.g. Krantz and Tversky, p. 156). This is because, for example, the additive model can be obtained by logarithmic transformation of the multiplicative model (having scale values on Re^+).

Now consider a product set $A_1 \times A_2 \times A_3$, where the values of each of three independent variables A_1 , A_2 , and A_3 , belong to non-empty and possibly infinite sets. For instance let $A_1 = \{a, b, c, \dots\}$, $A_2 = \{p, q, r, \dots\}$, and $A_3 = \{u, v, w, \dots\}$. Also, let a triple such as ‘*apu*’ denote an element of $A_1 \times A_2 \times A_3$ corresponding to level a of variable A_1 , level p of variable A_2 , and level u of variable A_3 . An empirical data structure consists of a finite number of levels n_1 (respectively, n_2 , n_3) of A_1 (respectively, A_2 , A_3), where it may be assumed that the dependent variable is represented by a weak order \leq on the $n_1n_2n_3$ elements. Hence, \leq denotes the *data* weak order.

This study addresses the issue of testing whether the $n_1n_2n_3$ elements created from the n_1 , n_2 , and n_3 values satisfy the key necessary axioms for the additive, distributive, and dual-distributive composition rules of PCM. Although the current study focuses on testing axioms of composition rules with three independent variables, these axioms are relevant for any number of independent variables (Krantz, 1968).

The *additive composition rule* assumes the existence of real-valued functions φ_k on A_k , $k = 1, 2, 3$, such that:

$$apu \leq bqv \Leftrightarrow \varphi_1(a) + \varphi_2(p) + \varphi_3(u) \leq \varphi_1(b) + \varphi_2(q) + \varphi_3(v), \quad (1)$$

¹Therefore, we do not consider axioms related to *sign dependence* (KLST, p. 329), which occurs when at least one of the variables has a *variable sign*, i.e. its domain is not included wholly in Re^- or wholly in Re^+ .

for all $a, b \in A_1$, all $p, q \in A_2$, and all $u, v \in A_3$, where φ_1, φ_2 , and φ_3 are interval scales.

The *distributive composition rule*² assumes real-valued functions φ_k on A_k , for $k=1,2,3$, such that:

$$apu \preceq bqv \iff [\varphi_1(a) + \varphi_2(p)]\varphi_3(u) \leq [\varphi_1(b) + \varphi_2(q)]\varphi_3(v), \tag{2}$$

for all $a, b \in A_1$, all $p, q \in A_2$, and all $u, v \in A_3$, where $[\varphi_1 + \varphi_2] \varphi_3$ is a ratio scale.

The *dual-distributive composition rule*² assumes real-valued functions φ_k on A_k , for $k=1, 2, 3$, such that:

$$apu \preceq bqv \iff [\varphi_1(a) \varphi_2(p)] + \varphi_3(u) \leq [\varphi_1(b) \varphi_2(q)] + \varphi_3(v) \tag{3}$$

for all $a, b \in A_1$, all $p, q \in A_2$, and all $u, v \in A_3$, where $\varphi_1 \varphi_2$ is a ratio scale.

Five PCM axioms provide a basis for testing whether the empirical order (data) supports any of the three polynomial composition rules. These five axioms are the following.

Axiom 1. (Single independence (I)) A_1 is *singly independent* of A_2 and A_3 if and only if (iff):

$$apu \preceq bpu \iff aqv \preceq bqv \tag{4}$$

is true for all $a, b \in A_1$, all $p, q \in A_2$, and all $u, v \in A_3$. The single independence of A_2 , and of A_3 , is similarly defined.

Axiom 2. (Double cancellation (DC)) *Double cancellation* holds for $A_1 \times A_2$ iff, with respect to every choice of $a, b, c \in A_1$ and $p, q, r \in A_2$, when:

$$aqu \preceq bru \text{ and } bpu \preceq cqu \text{ implies } apu \preceq cru \tag{5}$$

is true for all $u \in A_3$. Double cancellation on $A_1 \times A_3$, and on $A_2 \times A_3$, is similarly defined.

Axiom 3. (Joint independence (JI)) The pair $A_1 \times A_2$ is *jointly independent* iff:

$$apu \preceq bqu \iff apv \preceq bqv \tag{6}$$

is true for all $a, b \in A_1$, all $p, q \in A_2$, and all $u, v \in A_3$. The joint independence of the variable pair $A_1 \times A_3$, and of the pair $A_2 \times A_3$, is similarly defined. A structure that is jointly independent for all three variable pairs is denoted by JI3. Likewise, JI2 refers to two pairs satisfying joint independence, and JI1 refers to one pair.

Axiom 4. (Distributive cancellation) Distributive cancellation holds iff:

$$apu \preceq crv \text{ and } bqu \preceq dsv \text{ and } drv \preceq bpu \text{ implies } aqu \preceq csv \tag{7}$$

is true for all $a, b, c, d \in A_1$, all $p, q, r, s \in A_2$, and all $u, v \in A_3$.

²In Eqs. (2) and (3), variables 1 and 3, and 2 and 3, are interchangeable, as mentioned earlier. Therefore, the axiomatic framework, to be described, is the same no matter which permutation of the variables is considered.

Axiom 5. (Dual-distributive cancellation) Dual-distributive cancellation holds iff:

$$\begin{aligned}
 &csv \preceq arw \quad \text{and} \quad dpu \preceq btx \quad \text{and} \quad drx \preceq esu \quad \text{and} \quad eqy \preceq cty \quad \text{and} \\
 &aty \preceq dqy \quad \text{implies} \quad apv \preceq bqw
 \end{aligned}
 \tag{8}$$

is true for all $a, b, c, d, e \in A_1$, all $p, q, r, s, t \in A_2$, and $u, v, w, x, y \in A_3$.

If a structure violates any single independence or double cancellation test, or all three tests of joint independence, then the structure cannot (KLST, p. 345) be represented by any simple polynomial composition rule (for example, additive, distributive, or dual distributive).

A structure that passes all three single independence tests, all three double cancellation tests, and all three joint independence tests, is denoted by $I \wedge DC \wedge JI3$. The condition $I \wedge DC \wedge JI3$ is a necessary, but not sufficient, condition for the additive composition rule (Ad) to hold (e.g. Krantz and Tversky, 1971, p. 153), thus, Ad implies $I \wedge DC \wedge JI3$, but not vice-versa. Scott (1964) gives sufficient conditions for additivity when $n_1 n_2 n_3$ is finite; these conditions have been implemented in Sherman's (1977) additivity-test algorithm (applicable for conjoint structures).

Now consider a non-additive structure that satisfies $I \wedge DC \wedge JIh$, for some $h \in \{1, 2, 3\}$. This structure is consistent with the distributive rule if it satisfies distributive cancellation, and consistent with the dual-distributive rule if it satisfies dual-distributive cancellation. KLST (1971, pp. 343–344) prove that both distributive cancellation and dual-distributive cancellation are necessary conditions for an additive representation to be possible. Also a distributive representation need not satisfy dual-distributive cancellation, and a dual-distributive representation need not satisfy distributive cancellation (KLST, 1971, pp. 343–344).

The following illustrates earlier applications of the axiomatic approach. In the study by Ullrich and Painter (1974), the entries of a $3 \times 3 \times 2$ structure represented a subject's rank-order preference of $n_1 n_2 n_3 = 18$ job applicants. Each applicant had a unique combination of IQ level (one of $A_1 = \{\text{high, medium, low}\}$), experience level (one of $A_2 = \{5 \text{ years, 1 year, no past experience}\}$), and motivation (one of $A_3 = \{\text{low, high}\}$). The data of each subject was represented by a PCM composition rule. To give another example, each cell of a structure may indicate the binary-choice proportion for a group of subjects. A typical triple apu can then represent the relative frequency that subjects of group a successfully complete task p , as perceived by judge u . The $n_1 n_2 n_3$ proportion values can then be rank-ordered by their relative frequency prior to axiom testing. In this case, obviously, the selected composition rule (if one can be supported) represents the entire sample of subjects. KLST (1971, Chapter 7) and Krantz and Tversky (1971) mention several other research areas where conjoint composition rules are useful to measure and explain psychological variables, including behavior theory, decision theory, and expected utility theory. PCM has, in fact, been applied to data from tasks involving: decision-making under risk (e.g. Tversky, 1967; Coombs and Huang, 1970; Person and Barron, 1978), binaural perception (e.g. Levelt et al., 1972; Gigerenzer and Strube, 1983; Falmagne, 1976), contraceptive values (Nickerson and McClelland, 1988), attitude measurement (e.g. Michell, 1998), and signal detection (Scheiblechner, 1999).

It is well-known that because of their deterministic, algebraic form, PCM axioms, in general, cannot be ‘directly’ applied to social science data, which typically contain random and/or systematic error (e.g. Falmagne, 1976). However, Iverson and Falmagne (1985) show that the statistical ideas of order restricted inference (Robertson et al., 1988) provide a natural framework for testing (probabilistic) measurement theory axioms. Iverson and Falmagne’s method first specifies a model with inequality restrictions that correspond to those of a set of measurement axioms, and then employs a likelihood ratio test of the fit of these restrictions to data. They illustrate the usefulness of the method, but observe that several issues arise when interfacing statistics with axiomatic measurement theory.

Recently, Karabatsos (2001) and Karabatsos and Sheu (2001) demonstrate that Markov Chain Monte Carlo (MCMC) algorithms provide a method of order-restricted statistical inference that yields a simple way to test measurement axioms, and even to perform model selection between different axiomatic models. The MCMC methods seem to have eliminated most of the statistical issues, but one remains. That is, the order-restricted methods require the analyst to a-priori specify a single ordering on the data structure. However, as implied in the PCM axiom review above, and mentioned by Iverson and Falmagne (1985, p. 132), any given measurement axiom, or composition rule, is usually compatible with many possible orderings on the data structure, not just a single order. For instance, consider the task of testing whether the data of a 3×5 conjoint structure (statistically) agree with the axioms of single independence and double cancellation. To implement the statistical test with order restricted inference, the analyst needs to study, in turn, each of the 2773 possible orderings of a 3×5 structure (McClelland, 1977), and decide which, if any, satisfy both axioms.

This particular issue can be addressed by employing Monte Carlo methods to enumerate all possible (rank) orders of a conjoint structure. These methods can also facilitate a better understanding of the measurement axioms. Specifically, enumeration studies benefit the practice of axiomatic modeling in at least three ways. First, they demonstrate the restrictiveness of the measurement axioms. Second, they illustrate how the axioms characterize each of the composition rules. Third, they render available the set of possible data structures consistent with certain axioms, and certain composition rules. This third benefit provides information about the degree of structural similarity between the measurement axioms and between the different composition rules, and it may also provide a basis on which to develop new measurement axioms. Finally, the available set of orderings, coupled with order-restricted statistical inference, lead to focused statistical tests of the axioms, and aid in selecting the best-fitting axiomatic model. Returning to the example involving the 3×5 conjoint data structure, the analyst may explore the 2773 possible orderings available, and test the fit of the single ordering most ‘similar’ to the empirical ordering (data). The similarity measure is introduced later.

Three previous studies have enumerated various conjoint structures, although they did not address the issue of axiom testing. Arbuckle and Larimer (1976), who investigated structures with simple linear orders (rank orders with no ties), used Monte Carlo methods to determine the approximate number of two-way conjoint structures satisfying additivity. They also derived a formula that calculates the exact number of two-way

structures that satisfy single independence, this number being a function of the number of levels in each independent variable. Later, McClelland (1977) determined the exact number of 3×3 , 3×4 , 3×5 , and 4×4 structures that satisfy single independence (respectively, double-cancellation, additivity). Ullrich and Wilson (1993) extended McClelland's study and determined the exact number of 3×6 , 4×5 , 4×6 , 5×5 , $3 \times 3 \times 2$, $3 \times 4 \times 2$, and $3 \times 3 \times 3$ structures that satisfy the axioms required for an additive representation to exist. These authors found, as did McClelland (1977) and Arbuckle and Larimer (1976), that the number of structures satisfying the axioms is extremely small relative to the number of possible structures. Furthermore, the possibility of satisfying additivity sharply decreases as the number of items in the structure increases.

2. Research objectives

The present study has two general objectives. First, while Ullrich and Wilson (1993) have generated all $3 \times 3 \times 2$ structures satisfying the axioms required for an additive representation to exist, distributive rule (respectively, dual-distributive rule) structures have never been enumerated. Therefore, the Monte Carlo method is used to generate two sets of conjoint structures. The first set is compatible with the distributive composition rule (yielding the *distributive rule set*), and the second set is compatible with the dual-distributive composition rule (yielding the *dual-distributive rule set*). The next section describes the enumeration study. It illustrates how the axioms of joint independence, distributive cancellation, and dual-distributive cancellation, together, come to characterizing the distributive (respectively, dual-distributive) composition rule.

The section that follows describes a PCM analysis of a $3 \times 3 \times 2$ data set. This analysis demonstrates how Monte Carlo generated sets of conjoint structures are useful for statistics-based axiom testing, and for selecting the best fitting among the additive, distributive, or dual-distributive (axiomatically motivated) models. The demonstration employs the statistical framework developed by Karabatsos (2001). This is introduced in the following section.

3. Enumeration study

This enumeration study has five aims. First, to determine the restrictiveness of each of the three composition rules, i.e. to find out how much the possible data sets are limited by the requirement that each satisfies one (or more) of the composition rules. Ullrich and Wilson (1993) suggest that there are far more structures that are distributive or dual-distributive than there are structures that are additive. If such is the case, it implies that the restrictions of the distributive or dual-distributive composition rule are much weaker than those placed by the additive composition rule. Second, the study explores the conditions under which distributive cancellation holds for the dual-distributive set, and the conditions under which dual-distributive cancellation holds for the distributive set. Third, the study examines the number of joint independence tests (indicated by JI1,

J12, or J13) that the distributive set (respectively, the dual-distributive set) tends to satisfy. Fourth, in order to determine the structural proximity between the three PCM composition rules, the study measures the similarity between additive rule structures and the distributive rule set (respectively, with the dual-distributive rule set). Kendall's τ statistic, $\tau \in [-1, 1]$, provides a natural measure of similarity between two rank-ordered structures. The correlation measure τ is linearly related to Q , the minimum number of pair reversals necessary to transform the ranking of any one structure to the ranking of the other structure.

The present enumeration study only considers simple linear orderings on the $3 \times 3 \times 2$ conjoint structures. Current computer technology renders it difficult to enumerate conjoint structures for sets larger than $3 \times 3 \times 2$, even for simple linear orderings (see Ullrich and Wilson, 1993). However, a conjoint structure of this size can be used to distinguish between the additive, distributive, and dual-distributive composition rules (see KLST, 1971, Chapter 7).

3.1. Enumeration methods

The Monte Carlo generation of a single random $3 \times 3 \times 2$ structure consists of three-steps (e.g. Park and Miller, 1988). Step one generates three random numbers over $[0,1]$, orders them, and then assigns them to $A_1 = \{a, b, c\}$. This process repeats for $A_2 = \{p, q, r\}$, and for $A_3 = \{u, v\}$. Step two generates a structure satisfying the distributive (respectively, dual-distributive) rule, yielding a $3 \times 3 \times 2$ matrix of real numbers. Specifically, to generate a distributive rule (respectively, dual-distributive rule) structure, the value of any one of the 18 cells, say apu , is determined by calculating $(\underline{a} + \underline{p})\underline{u}$ (respectively, $\underline{a}(\underline{p} + \underline{u})$), where \underline{a} is the random value assigned to element a in step one (\underline{p} and \underline{u} are similarly defined). The generated matrix of real numbers are then rank ordered $1, \dots, 18$ (if a tie occurs, the process returns to step one). In step three, the conjoint analysis program CPCJM (Ullrich and Wilson, 1990) determines which PCM tests are satisfied by the rank-ordered structure³. Since the distributive and dual-distributive functions, used to generate the structures, are simple polynomials, the generated structure belongs to one of four classes:

1. $I \wedge DC \wedge J11$
2. $I \wedge DC \wedge J12$
3. $I \wedge DC \wedge J13 \wedge \neg Ad$
4. Ad.

The above Monte Carlo procedure is used to generate billions of random structures satisfying the distributive rule (respectively, the dual-distributive rule). For each set so generated, the procedure discards duplicate structures. Each of these two sets is then compared with the entire set of $3 \times 3 \times 2$ structures enumerated by Ullrich and Wilson

³CPCJM implements both Krantz and Tversky's (1971) conjoint measurement tests and Sherman's (1977) additivity algorithm.

(1993) to determine the degree of intersection between the three sets. Those authors identified all 3972 additive rule structures of the $3 \times 3 \times 2$ set.

The similarity of a particular distributive (respectively, dual-distributive) structure i to an additive one is measured by using Kendall's τ rank-order correlation coefficient between that particular structure and each of the 3792 additive structures in turn. The maximum τ correlation value over the 3792 correlated pairs, denoted $\tau(\text{Ad})_i$, measures the proximity between that distributive (respectively, dual-distributive) structure i and structures satisfying the additive composition rule. The same correlation procedure is repeated for each and every distributive (respectively, dual-distributive) structure. Hence, the correlation procedure for the distributive set generates a distribution with values $\tau(\text{Ad})_m$, $m = 1, \dots, M$, where M is the number of distributive structures. Similarly, the procedure for the dual-distributive set generates a second distribution with values $\tau(\text{Ad})_g$, $g = 1, \dots, G$, where G is the number of dual-distributive structures. These two distributions convey the proximity between the additive, distributive, and dual-distributive composition rules of PCM.

3.2. Enumeration results

For the distributive rule set (respectively, dual-distributive rule set), eight billion random matrices were incrementally generated, two billion at a time, while discarding duplicate structures. The procedure generated $M = 27\,382$ distinct distributive rule structures, and $G = 31\,996$ distinct dual-distributive rule structures. The final two billion random matrices added less than 100 new matrices to the distributive rule set (respectively, dual-distributive rule set).

The resulting counts of the two sets strongly suggest that the requirements of the distributive composition rule are stricter than the requirements of the dual-distributive composition rule, with the additive composition rule (having 3792 structures in the above simulations) being the strictest by far. The full set of $3 \times 3 \times 2$ additive, distributive, dual-distributive structures are available from the authors.

Column 1 of Table 1 lists the total number of $3 \times 3 \times 2$ structures, either additive or non-additive, satisfying single independence, double cancellation, and one or more joint independence axioms. The results of this column were determined by a previous Monte Carlo study (Ullrich and Wilson, 1993). The second and third columns detail the same results for the distributive set and the dual-distributive set, and show that all the 3792 additive structures satisfy distributive and dual-distributive cancellation, as expected.

Note that the distributive set and the dual-distributive set overlap only for the 3792 additive structures. The set of $(= 27\,382 - 3792)$ 23 590 non-additive distributive rule structures does not intersect the set of 28 204 $(= 31\,996 - 3792)$ non-additive dual-distributive rule structures. Thus, in the case $I \wedge DC \wedge (JI1 \vee JI2 \vee JI3)$, if a $3 \times 3 \times 2$ structure violates additivity and satisfies distributive cancellation, then it violates dual-distributive cancellation; if a $3 \times 3 \times 2$ structure violates additivity and satisfies dual-distributive cancellation, then it violates distributive cancellation. There is a small but non-zero possibility that counterexamples to this generalization exist among structures that were not generated by the Monte Carlo procedure. This seems unlikely, given that eight billion random matrices were generated for each set. Hence, a future

Table 1
The number of $3 \times 3 \times 2$ structures satisfying conjoint measurement axioms^a

Condition	Set of $3 \times 3 \times 2$ structures satisfying the given condition		
	Total set of structures ^b	Structures within the distributive set	Structures within the dual-distributive set
$I \wedge DC \wedge JI1$	124 912	12 759	16 397
$I \wedge DC \wedge JI2$	45 228	10 397	11 335
$I \wedge DC \wedge JI3 \wedge \neg Ad$	1100	434	472
Ad	3792	3792	3792
Total number of structures: additive and non-additive	175 032	27 382	31 996
Total number of non-additive structures satisfying composition rule		Distributive 23 590 (27 382–3792)	Dual-distributive 28 204 (31 996–3792)

^a The total number of possible linear rank orderings on a $3 \times 3 \times 2$ structure is 6.4×10^{15} (Ullrich and Wilson, 1993).

^b This column contains the $3 \times 3 \times 2$ structures enumerated by Ullrich and Wilson (1993).

study might attempt to find an analytical proof generalizing these results for all $3 \times 3 \times 2$ conjoint structures to conjoint structures of any size. Such a proof has yet to be derived.

Table 2 represents the enumeration results of Table 1 as relative frequencies. The top section of Table 2 refers to the distributive and dual-distributive sets containing both additive and non-additive structures. About 50% of the structures in each set pass one

Table 2
Proportions of structures satisfying PCM axioms (exact to three significant digits)

Additive and non-additive structures		
	Distributive set	Dual-distributive set
$I \wedge DC \wedge JI1$	0.466	0.512
$I \wedge DC \wedge JI2$	0.380	0.354
$I \wedge DC \wedge JI3 \wedge \neg Ad$	0.016	0.015
Ad	0.138	0.119
Total number of structures: Additive and non-additive	27 382	31 996
Non-additive structures		
	Distributive rule set	Dual-distributive rule set
$I \wedge DC \wedge JI1$	0.541	0.581
$I \wedge DC \wedge JI2$	0.441	0.402
$I \wedge DC \wedge JI3 \wedge \neg Ad$	0.018	0.017
Total number of non-additive structures	23 590	28 204

joint independence test, where the dual-distributive set more frequently satisfies JI1. Also, in each set, less than 40% satisfy JI2, where the distributive set possesses this condition more frequently. Less than 15% of the structures is additive (14% of the distributive set, 12% of the dual-distributive set). Less than 2% of structures are both non-additive and pass all three joint-independence tests (1.6% of the distributive set, 1.5% of the dual-distributive set). In fact, of all JI3 structures, almost 90% is additive (89% of $I \wedge DC \wedge JI3$ structures in the distributive set is additive; 88% of $I \wedge DC \wedge JI3$ structures in the dual-distributive set is additive). The bottom section of Table 2 refers to non-additive structures, either characterized by the distributive composition rule or the dual-distributive composition rule. That section shows a pattern of results within the distributive rule set (respectively, within the dual-distributive rule set) that is similar to that already discussed for additive and non-additive structures taken together. Namely, the distributive rule set tends to satisfy more joint independence tests than does the dual-distributive rule set. Also, only about 2% of structures of either composition rule satisfy JI3.

Table 3 provides the two $\tau(\text{Ad})$ distributions, based on the correlation of each of the 27 382 structures of the distributive set, and each of the 31 996 structures of the dual-distributive set, with additive structures. The two rank-order distributions indicate that the additive, distributive, and dual-distributive structures are very similar, with the distributive structures being slightly more similar to additivity than are the dual-distributive structures. Of the (maximum possible) $153 (= \frac{1}{2}18(18-1))$ pair reversals to

Table 3
Correlation distributions that illustrate the proximity between the additive rule, distributive rule, and dual-distributive rule, for $3 \times 3 \times 2$ conjoint structures

$\tau(\text{Ad})$	Q	Distributive set		Dual-distributive set	
		Count	$\text{Pr}[\leq Q]$	Count	$\text{Pr}[\leq Q]$
1.000	0	3792	0.138	3792	0.119
0.993	1	5538	0.341	5821	0.300
0.987	2	5707	0.549	6247	0.496
0.980	3	4585	0.717	5372	0.664
0.974	4	3154	0.832	3893	0.785
0.967	5	2077	0.908	2744	0.871
0.961	6	1338	0.957	1826	0.928
0.954	7	714	0.983	1137	0.964
0.948	8	325	0.994	638	0.984
0.941	9	110	0.998	326	0.994
0.935	10	34	0.9997	136	0.998
0.928	11	8	1.000	48	0.9995
0.922	12	0	1.000	12	0.9999
0.915	13	0	1.000	4	1.000
Total	27 382		31 996		

A single value of $\tau(\text{Ad})$ is the Kendall's τ correlation of a single distributive (dual-distributive) structure to additivity. It is obtained by correlating that structure to each and every one of the 3792 additive structures, and then finding the maximum τ over the 3792 correlated pairs. The measure Q is the minimum number of pair reversals necessary to transform a distributive (dual-distributive) structure into an additive one. The correlation $\tau(\text{Ad})$ is a linear function of Q .

convert a given structure into an additive one, at most only 11 are needed for a distributive structure and 13 for a dual-distributive structure (though the maximum may not be achievable). In fact, approximately 95% of the distributive structures need at most only six pair reversals; $Q \leq 6$ corresponds to a τ correlation range of 0.961 and 1. About 95% of the dual-distributive structures need at most only seven pair reversals; $Q \leq 7$ corresponds to a τ correlation range of 0.954 and 1.

Of course, the number of possible structures rapidly increases as a function of conjoint structure size (e.g. Ullrich and Wilson, 1993). Therefore, the enumeration of larger conjoint structures in future studies may reveal that the additive, distributive, and dual-distributive sets are more dissimilar to each other, and also that the non-additive distributive set fails dual-distributive cancellation, and the non-additive dual-distributive set fails distributive cancellation. Also, given the results of Table 2, it seems that among these larger structures, those in the distributive set will be more successful in passing the joint independence tests than those in the dual-distributive set.

4. PCM axiom testing: an empirical example

This section uses a data set obtained from a signal detection experiment (Schwarz, 1990), to demonstrate how the conjoint structures generated for that data set aid in statistical axiom testing and in selecting the best-fitting model (additive, distributive, or dual-distributive), each motivated by a particular set of axioms. First, it is necessary to describe the data set, and to present the statistical framework (Karabatsos, 2001) used for data analysis.

4.1. Experiment and data

A data set was obtained from a signal detection experiment performed by Schwarz (1990) (analyzed under the two-factor additive conjoint model in Scheiblechner, 1999). On a single trial, the subject was randomly exposed to one of three stimulus levels (S_0 , S_2 , S_5). The ‘standard’ stimulus S_0 was a rectangle of 2560 (64×40) pixels on a monochromatic monitor, of which 850 were ‘darkened’ randomly. The ‘test’ stimuli S_2 (respectively, S_5) had 890 (respectively, 950 pixels). For each trial, the participant’s task was to indicate whether the presented stimulus was the standard stimulus (S_0), by giving a ‘No’ response, or a test stimulus (either S_2 or S_5), by giving a ‘Yes’ response. Hence, ‘No’ is the correct response to a standard stimulus, and ‘Yes’ is the correct response to a test stimulus.

The experiment also varied the level of payoff to each subject. For each of the three payoff-conditions C1, C2, and C3, each subject was exposed to 1080 S_0 , 216 S_2 , and 216 S_5 presentations. In the strict-payoff condition C1, a correct response to a standard stimulus (‘No’) yielded 3 points for the subject, and a correct response to a test stimulus (‘Yes’) yielded 1 point. An incorrect response a standard stimulus (‘Yes’) cost 3 points, and an incorrect response to a test stimulus (‘No’) cost 1 point. In medium-payoff condition C2, a correct response yielded 2 points, and an incorrect response cost 2 points. In lax-payoff condition C3, a correct response to a standard stimulus yielded 1

Table 4
Proportions of ‘yes’ responses of two subjects (B and C), by pay-off level (C1, C2, C3), and stimulus level (S₀, S₂, S₅)

Stimulus	Subject B			Subject C		
	Pay-off level			Pay-off level		
	C1	C2	C3	C1	C2	C3
S ₀	0.168	0.296	0.433	0.219	0.441	0.449
S ₂	0.375	0.644	0.718	0.495	0.653	0.773
S ₅	0.856	0.926	0.991	0.852	0.861	0.986

Total proportion by:
 Subject Subject B: 0.4280. Subject C: 0.484
 Each is based on 4536 observations
 Pay-off level: C1: 0.322. C2: 0.484. C3: 0.563
 Each is based on 3024 observations
 Stimulus S₀: 0.334 (6480 observations)
 S₂: 0.610. S₅: 0.912 (each has 1296 observations)

Each of the six cells in level S₀ contains 1080 presentations of the stimulus. Each of the six cells in level S₂, and level S₅, contains 216 presentations of the stimulus.

point, and a correct response to a test stimulus yielded 3 points. An incorrect response cost 1 point for a standard stimulus, and yielded 3 points for the test stimulus.

The 3×3×2 data set of Table 4 contains the proportion of ‘Yes’ responses for each of the three stimuli, and the three payoff levels for each of two subjects. As one might expect, the proportions of ‘Yes’ responses generally increase as a function of stimulus level (increasing number of pixels), and as the pay-off conditions become less restrictive. Also, Subject C has a greater tendency than does Subject B to respond ‘Yes.’

4.2. A framework for probalistic measurement theory

As shown in Table 4, the 3×3×2 data set is represented by the 18 proportions:

$$\mathbf{p} = (p_j | p_j = n_j/N_j; j = 1, \dots, 18) \in [0,1]^{18} \tag{9}$$

where n_j is the number of ‘Yes’ responses in N_j trials. Correspondingly (e.g. Karabatsos, 2001),

$$\Theta = (\theta_j | j = 1, \dots, 18) \in S \subseteq [0,1]^{18} \tag{10}$$

where S is a proper subset of [0,1]¹⁸, represents the expected proportion of ‘Yes’ responses predicted by an axiomatic model (additive, distributive, or dual distributive). Thus the values in Eq. (10) satisfy the inequality restrictions assumed by the relevant composition rule.

The statistical framework to be applied treats the parameters of Θ as random, and therefore adopts a Bayesian perspective, which states the posterior distribution of Θ, conditional on the responses $\mathbf{p}=(p_1, \dots, p_j, \dots, p_{18})$, by:

$$\pi(\Theta|\mathbf{p}) = \frac{L(\Theta|\mathbf{p})\pi(\Theta)}{\int L(\Theta|\mathbf{p})\pi(\Theta)d\Theta} \tag{11}$$

where the denominator (integral) is the marginal density. The likelihood L is given by the binomial distribution:

$$L(\mathbf{p}|\Theta) = \prod_{j=1}^{18} \binom{N_j}{n_j} \theta_j^{n_j} (1 - \theta_j)^{N_j - n_j}, \tag{12}$$

which assumes that the subjects’ responses are independent, conditional on Θ . The inequality restrictions of the axiomatic model are specified by the prior distribution $\pi(\Theta)$ that constrains the order of the 18 empirical proportions. The prior $\pi(\Theta)$ is derived by specifying independent uniform prior distributions over the elements of Θ , i.e.

$$\pi(\theta_j) = \begin{cases} [\max(\theta_j) - \min(\theta_j)]^{-1} & \text{iff } \min(\theta_j) \leq \theta_j \leq \max(\theta_j) \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

where the inequality constraints force appropriate values to exist with $0 \leq \min(\theta_j) < \max(\theta_j) \leq 1$.

Karabatsos (2001) describes a general MCMC algorithm that can generate samples from the posterior $\pi(\Theta|\mathbf{p})$, without requiring calculation of the complicated integral (the marginal density in Eq. (11)). The details of this algorithm are presented in Appendix A.

After generating samples from the posterior $\pi(\Theta|\mathbf{p})$, the measurement axioms of a model are tested by comparing the elements of the data structure $\mathbf{p} = (p_1, \dots, p_{18})$ to the posteriors of the corresponding model parameter elements $\Theta = (\theta_1, \dots, \theta_{18})$. For each and every j , this is done by checking whether p_j is contained within the 0.25 and 0.975 quantiles of the generated posterior distribution of θ_j . These quantiles define the 95% posterior interval of θ_j .

For each of three axiomatic models, the MCMC algorithm was applied to analyze the data set. One set of parameters Θ was estimated under the order restrictions of the additive composition rule, a second set assumed the distributive rule restrictions, and the third set employed the constraints of the dual-distributive rule.

Using the same MCMC algorithm, the Deviance Information Criterion (DIC, Spiegelhalter et al., 2001) can be computed for each model (additive, distributive, and dual-distributive), in order to perform model selection. The model with the lowest DIC value is identified as best fitting, i.e. having the best tradeoff between underfit and overfit. For the signal-detection data set, a model’s DIC value is based on the deviance function:

$$D(\Theta) = 2 \sum_{j=1}^{18} \left[\binom{n_j}{n_j} \ln\left(\frac{n_j}{N_j \theta_j}\right) + (N_j - n_j) \ln\left(\frac{N_j - n_j}{N_j - N_j \theta_j}\right) \right] \tag{14}$$

(e.g. McCullagh and Nelder, 1983, p. 81), which measures the ‘distance’ of the model from the empirical data set (to prevent division by zero in Eq. (14), any θ_j value of 0 is

set to 0.00001, and any value of 1 is set to 0.99999). For a single axiomatic model, the DIC is determined by:

$$DIC = D(\hat{\Theta}) + 2(\overline{D(\Theta)} - D(\hat{\Theta})) \tag{15}$$

where the first term is the deviance (15) evaluated at the posterior mean $\hat{\Theta} = (\hat{\theta}_1, \dots, \hat{\theta}_{18})$, and the second term penalizes model complexity, where $\overline{D(\Theta)}$ is the posterior mean of the deviance. The penalty term represents the ‘effective’ number of model parameters, and measures the sensitivity of the parameters to changes in the data (see Spiegelhalter et al., 2001). Unlike the majority of model selection methods (see Myung et al., 2000, *Journal of Mathematical Psychology*, Special Issue on Model Selection), DIC is able to measure the complexity of qualitative models, such as the three composition rule models of PCM.

4.3. Data analysis

This sub-section shows that having the PCM axioms allows a detailed analysis of the data structure. The axioms also provide a basis on which to determine the most likely composition rule (additive, distributive, or dual-distributive) for the data structure having three stimulus levels ($A_1 = \{S_0, S_2, S_5\}$), three levels of payoff ($A_2 = \{C1, C2, C3\}$), and two subjects ($A_3 = \{B, C\}$).

Table 5a gives the empirical ranks (from smallest to largest) of the 18 observed

Table 5
Rank-order representations of Table 4

Stimulus	Subject B			Subject C		
	Pay-off level			Pay-off level		
	C1	C2	C3	C1	C2	C3
(a) Empirical ranks						
S_0	1	3	5	2	6	7
S_2	4	9	11	8	10	12
S_5	14	16	18	13	15	17
(b) Ranks of the additive composition rule						
S_0	1	3	6	2	5	8
S_2	4	9	11	7	10	12
S_5	13	15	17	14	16	18
(c) Ranks of the distributive composition rule						
S_0	1	3	6	2	5	8
S_2	4	9	11	7	10	12
S_5	13	15	16	14	17	18
(d) Ranks of the dual-distributive composition rule						
S_0	1	3	5	2	6	8
S_2	4	9	11	7	10	12
S_5	13	15	17	14	16	18

proportions of Table 4. This order corresponds to $\neg I \wedge DC \wedge JI1$ (as determined by the program CPCJM). The ‘apparent’ single independence violations of A_3 (indicated by $\neg I$) are due to the data pairs $[(S_5, C1, C)=13] < [(S_5, C1, B)=14]$, $[(S_5, C2, C)=15] < [(S_5, C2, B)=16]$, and $[(S_5, C3, C)=17] < [(S_5, C3, B)=18]$. On the other hand, there are two apparent joint independence violations, one on $A_1 \times A_2$ due to the data pair $[(S_0, C3, C)=7] < [(S_2, C1, C)=8]$, and the other on $A_2 \times A_3$ due to the data pair $[(S_0, C3, B)=5] < [(S_0, C2, C)=6]$. Of course, all of the mentioned violations are only suggestive, since it is possible that they are only a result of random error, instead of a systematic underlying incompatibility with the axioms. This is a reasonable perspective, considering that each of the mentioned pairs only slightly violates the axioms (and this violation only being in terms of ranks). Therefore, the MCMC statistical procedure is applied in order to decide whether the empirical ordering, determined by the proportions of Table 4, satisfies the order restrictions determined by the measurement axioms.

As mentioned earlier, the analyst faces the statistical issue of having to select from a very large set of possible rank orders in order to perform axiom and composition rule testing. Now, the generated additive rule, distributive rule, and dual-distributive rule structures are already at hand. The ranks in Table 5b represent a single additive structure ($I \wedge DC \wedge JI3 \wedge Ad$) from the available 3792 $3 \times 3 \times 2$ additive structures. Of all the 3792 additive structures, the additive structure in Table 5b has the highest rank correlation ($\tau=0.935$) with the empirical ranks of Table 5a. Note that Table 5b has the reverse of the empirical order for five pairs: each of the three pairs violating A_3 independence, the pair violating $A_1 \times A_2$ joint independence, and the pair violating $A_2 \times A_3$ joint independence. The ranks in Table 5c represent the most similar distributive rule structure ($I \wedge DC \wedge JI2$) among the 23 590 (non-additive) distributive rule structures. This structure has the highest rank correlation ($\tau=0.922$) of such distributive structures with the empirical rank ordering. Table 5c has the reverse of the empirical order for the five pairs already mentioned, but with a different $A_2 \times A_3$ joint independence violation due to the data pair $[(S_5, C3, B)=16] < [(S_5, C2, C)=17]$. Finally, of the 28 204 (non-additive) dual-distributive rule structures, Table 5d gives the ranks of the structure ($I \wedge DC \wedge JI2$); these ranks are the most similar to the empirical ranks ($\tau=0.948$). Table 5d reverses the empirical order of each of the three pairs violating A_3 independence, and of the pair violating $A_1 \times A_2$ joint independence, but not of the pair violating $A_2 \times A_3$ joint independence.

For each of the three composition rule models, the MCMC algorithm generated 4500 samples from the posterior $\pi(\theta|\mathbf{p})$ (after discarding the initial 500 samples generated by the algorithm). The additive model specified a prior distribution $\pi(\theta)$ using the ranks in Table 5b as the order restrictions. The distributive (respectively, dual-distributive) model provide the prior order restrictions used for Table 5c (respectively, for Table 5d). The DIC concluded that the additive composition rule model (DIC=35.1, $D(\bar{\theta})=12.4$, penalty=22.7) and the dual-distributive composition rule model (DIC=35.0, $D(\bar{\theta})=12.3$, penalty=22.7) have about the same data fit, while the distributive composition rule is decisively the worst fitting model (DIC=61.7, $D(\bar{\theta})=38.4$, penalty=23.3).

For the additive composition rule model, Table 6 gives, for each of the 18 cells, the posterior mean $\bar{\theta}_j$, and the 95% posterior interval of θ_j . Each of the three underlined

Table 6

MCMC posterior estimates of the ‘yes’ response proportions, assuming the order restrictions of the additive composition rule (Table 5b)

Stimulus	Subject B			Subject C		
	Pay-off level			Pay-off level		
	C1	C2	C3	C1	C2	C3
S_0	<i>0.168</i>	<i>0.296</i>	<i>0.439</i>	<i>0.220</i>	<i>0.427</i>	<i>0.467</i>
	0.148, 0.189	0.270, 0.325	0.417, 0.459	0.196, 0.247	0.402, 0.449	0.441, 0.491
S_2	<i>0.373</i>	<i>0.624</i>	<i>0.717</i>	<i>0.455</i>	<i>0.661</i>	<i>0.774</i>
	0.316, 0.423	0.567, 0.673	0.666, 0.766	<u>0.431, 0.482^b</u>	0.612, 0.710	0.724, 0.820
S_5	<i>0.836</i>	<i>0.887</i>	<i>0.980</i>	<i>0.859</i>	<i>0.897</i>	<i>0.988</i>
	0.796, 0.868	<u>0.862, 0.914^a</u>	0.963, 0.991	0.826, 0.887	<u>0.870, 0.923^c</u>	0.975, 0.996

In each cell, the number in *italics* refers to the posterior mean, and the number pair is the lower and upper bounds of the 95% posterior interval (based on 0.025 and 0.975 posterior quantiles, respectively). Each of the superscripts ‘a’, ‘b’, and ‘c’ indicate a cell that contributes to axiom violations of the additive model, where the observed proportion (given in Table 4) falls outside the 95% posterior interval of that model. The observed proportion referred to by ‘a’ is 0.926; ‘b’ refers to 0.495, and ‘c’ pertains to 0.861.

intervals in Table 6 indicates a cell where the observed proportion (of Table 4) lies outside the 95% interval. Though the posteriors of the dual-distributive model are not presented, the same three observed proportions are outliers with respect to that model.

Two of the underlined cells of Table 6 refer to the empirical order $[(S_5, C2, C)=15] < [(S_5, C2, B)=16]$ that violates A_3 single independence. The third underlined cell of Table 6 refers to the empirical ordering $[(S_0, C3, C)=7] < [(S_2, C1, C)=8]$ that violates $A_1 \times A_2$ joint independence. These violations must be attributed to a systematic underlying incompatibility with the additive model, not only to random error.

As mentioned already, the empirical orders $[(S_5, C1, C)=13] < [(S_5, C1, B)=14]$ and $[(S_5, C3, C)=17] < [(S_5, C3, B)=18]$ suggest violations of A_3 single independence, and the empirical ordering $[(S_0, C3, B)=5] < [(S_0, C2, C)=6]$ suggests a violation of $A_2 \times A_3$ joint independence. However, the posterior intervals of each of the six cells, in the three pairs, contain the corresponding observed proportions of Table 4. Therefore, these order violations are only due to random error; they do not indicate a systematic underlying incompatibility with the additive model.

In summary, the signal-detection data of Table 4 can, in general, be described by an additive composition rule, because each of the three violating cell proportions (0.926, 0.495, 0.861) is close to each of the relevant 95% posterior intervals ($[0.862, 0.914]$, $[0.431, 0.482]$, $[0.870, 0.923]$) for an additive model.

5. Conclusion and discussion

In their review of representational measurement theory, Luce and Narens (1994)) describe 15 unsolved problems, one of them being:

Problem 2. *Specify a probabilistic version of measurement theory and the related statistical methods for evaluating whether or not a data set supports or refutes specific measurement axioms (Luce and Narens, 1994, p. 227).*

The present study introduced and applied practical Monte Carlo methods to address this particular problem. Specifically, for a given axiomatic model, measurement structures were generated by Monte Carlo methods in order to find a single (rank) order restriction to represent the prior distribution $\pi(\Theta)$, and then Markov Chain Monte Carlo was implemented to generate a sample from the model's posterior $\pi(\Theta|\mathbf{p})$, assuming that prior. This study showed that, given a generated sample from the posterior, it is possible to test axioms in detail, and to select the axiomatic model that best fits the data.

As mentioned earlier, generating measurement structures using Monte Carlo methods is computationally expensive. Therefore, we intend in our to develop a deterministic optimization routine that, given the data set, finds the most likely rank order under the assumed PCM composition rule (model). The routine would provide an alternative and faster method for deriving the prior distribution $\pi(\Theta)$ for a given axiomatic model.

Finally, the current study selected DIC as the model selection criterion, only because it offers an easy way to compute the complexity of qualitative models through MCMC. The DIC criterion may not necessarily be the best available for axiomatic models. Therefore, future research will also investigate the quality of DIC for axiomatic models, and compare this quality to other model selection criteria applicable to qualitative models. Examples of other criteria include posterior predictive model selection (Laud and Ibrahim, 1995; Gelfand and Ghosh, 1998) and the minimum description length method (e.g. Myung, 2000).

Acknowledgements

The authors give special thanks to the guest editor A.A.J. Marley and three anonymous referees for their helpful and detailed comments on earlier versions of this article. Portions of this study were presented at the 34th Annual Meeting of the Society for Mathematical Psychology, Brown University, Providence, Rhode Island. The research was supported, in part, by Spencer Foundation Grant SG200100020, George Karabatsos, Principal Investigator.

Appendix A. MCMC Algorithm

For a particular axiomatic model chosen for data analysis, the following presents the technical details of the MCMC algorithm, to illustrate how the Θ posterior and DIC are estimated. The algorithm is described within the context of the signal detection data set.

First, starting values $\Theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_{18}^{(0)})$ need to initiate the MCMC algorithm. These values are randomly chosen, under the condition that they follow the rank order restrictions specified by the axiomatic model through the prior (13). Now, let $\theta_{<j} =$

$\{\theta_1, \theta_2, \dots, \theta_{j-1}\}$ (where $\theta_{<j} = \{\emptyset\}$ when $j=1$), and let $\theta_{>j} = \{\theta_{j+1}, \theta_{j+2}, \dots, \theta_{18}\}$ (where $\theta_{>j} = \{\emptyset\}$ when $j=18$). For a single iteration t , $t=1, \dots, T$, the MCMC algorithm consists of six, looping steps:

Step 1. Set $j=0$.

Step 2. Let $j=j+1$ and draw a random number R_j from the uniform distribution on $[0, 1]$.

Step 3. Draw a candidate θ_j^C from the uniform distribution on $[\min(\theta_j)^{(t)}, \max(\theta_j)^{(t)}]$.

Step 4. Decide:

$$\theta_j^{(t)} = \begin{cases} \theta_j^C & \text{iff } R_j \leq \frac{L(\mathbf{p}|\theta_j^C, \theta_{<j}^{(t)}, \theta_{>j}^{(t-1)})}{L(\mathbf{p}|\theta_j^{(t-1)}, \theta_{<j}^{(t)}, \theta_{>j}^{(t-1)})} \\ \theta_j^{(t-1)} & \text{otherwise} \end{cases}$$

Step 5. If $j < 18$, repeat Steps 2, 3, 4, and 5. If $j = 18$, proceed to Step 6.

Step 6. If $j = 18$, iteration t ends. Compute the deviance for iteration t , $D(\Theta^{(t)})$, where $\Theta^{(t)} = (\theta_1^{(t)}, \dots, \theta_{18}^{(t)})$.

In Step 3, $\min(\theta_j)^{(t)}$ is the current value of a parameter specified (by the prior) to be one rank below the parameter θ_j . If the prior constrains θ_j to have the lowest value among $(\theta_j | 1 \leq j \leq 18)$, then $\min(\theta_j)^{(t)} = 0$ for all iterations t . Likewise, $\max(\theta_j)^{(t)}$ refers to the current value of the parameter specified to be one rank above θ_j . If the prior constrains θ_j to have the highest value among the 18 parameters, then $\max(\theta_j)^{(t)} = 1$ for all t .

This MCMC algorithm characterizes a hybrid Metropolis Hastings–Gibbs sampler; descriptions of such a sampler are found in many standard texts on MCMC (e.g. Carlin and Louis, 1998). The ratio in (A1) is a posterior ratio of the form $\pi(\Theta^C|\mathbf{p})/\pi(\Theta^{(t-1)}|\mathbf{p})$, simplified to a likelihood ratio. This is because the marginal density constant (the integral in Eq. (11)) and the prior density cancel out of the numerator and denominator. With respect to the prior, the value of $\pi(\theta_j^C)$ of the numerator and the value of $\pi(\theta_j^{(t)})$ in the denominator are equal to $[\max(\theta_j)^{(t)} - \min(\theta_j)^{(t)}]^{-1}$.

The analysis task is to repeat the MCMC algorithm for a ‘large’ number of iterations T , because as T approaches infinity, the iterates $\{\Theta^{(1)}, \Theta^{(2)}, \dots, \Theta^{(t)}, \dots, \Theta^{(T)}\}$ converge to a sample from the posterior $\pi(\Theta|\mathbf{p})$ (for the conditions of MCMC convergence, see Tierney, 1994). Therefore, the posterior mean $\bar{\Theta} = (\bar{\theta}_1, \dots, \bar{\theta}_{18})$ and the quantiles for the 95% posterior interval are estimated directly from $\{\Theta^{(1)}, \Theta^{(2)}, \dots, \Theta^{(t)}, \dots, \Theta^{(T)}\}$. Also, for DIC, the posterior mean of the deviance $\bar{D}(\Theta)$ is estimated by averaging over the sample $\{D(\Theta^{(1)}), D(\Theta^{(2)}), \dots, D(\Theta^{(t)}), \dots, D(\Theta^{(T)})\}$, which is generated by Step 6 of the algorithm.

Note that prior to calculating the estimates of the Θ posteriors and DIC, it is wise to first discard the first few hundred samples generated by the MCMC algorithm, as they may depend on the randomly selected starting values $\Theta^{(0)}$.

References

- Arbuckle, J., Larimer, J., 1976. The number of two-way tables satisfying certain additivity axioms. *Journal of Mathematical Psychology* 12, 89–100.

- Carlin, B.P., Louis, T.A., 1998. *Bayes and Empirical Bayes Methods For Data Analysis*. Chapman and Hall/CRC, Boca Raton, FL, first reprint.
- Coombs, C.H., Huang, L.C., 1970. Polynomial psychophysics of risk. *Journal of Mathematical Psychology* 7, 317–338.
- Falmagne, J.-C., 1976. Random conjoint measurement and loudness summation. *Psychological Review* 83, 65–79.
- Gigerenzer, G., Strube, G., 1983. Are there limits to binaural additivity of loudness? *Journal of Experimental Psychology: Human Perception and Performance* 9, 126–136.
- Gelfand, A.E., Ghosh, S.K., 1998. Model choice: A minimum posterior predictive loss approach. *Biometrika* 85, 1–13.
- Iverson, G., Falmagne, J.-C., 1985. Statistical issues in measurement. *Mathematical Social Sciences* 10, 131–153.
- Karabatsos, G., 2001. The Rasch model, additive conjoint measurement, and new models of probabilistic measurement theory. *Journal of Applied Measurement* 2, 389–423.
- Karabatsos, G., Sheu, C.-F., 2001. Testing measurement theory axioms with Markov Chain Monte Carlo. Paper presented at the 34th Annual Meeting of the Society for Mathematical Psychology, Brown University, Providence, Rhode Island.
- Krantz, D.H., 1968. A survey of measurement theory. In: Dantzig, G.B., Veinott, Jr. A.F. (Eds.), *Mathematics of the Decision Sciences*. American Mathematical Society, Providence, RI.
- Krantz, D.H., Luce, R.D., Suppes, P., Tversky, A., 1971. Foundations of measurement. In: *Additive and Polynomial Representations*. Academic Press, New York.
- Krantz, D.H., Tversky, A., 1971. Conjoint-measurement analysis of composition rules in psychology. *Psychological Review* 78, 151–169.
- Laud, P.W., Ibrahim, J.G., 1995. Predictive model selection. *Journal of the Royal Statistical Society, Series B* 57, 247–262.
- Luce, R.D., Narens, L., 1994. Fifteen problems concerning the representational theory of measurement. In: Humphreys, P. (Ed.), *Patrick Suppes: Scientific Philosopher*, Vol. 2, pp. 219–249.
- McClelland, G., 1977. A note on Arbuckle and Larimer. The number of two way tables satisfying certain additivity axioms. *Journal of Mathematical Psychology* 15, 292–295.
- McCullough, P., Nelder, J.A., 1983. *Generalized Linear Models*. Chapman and Hall, London.
- Michell, J., 1998. Sensitivity of preferences and ratings to ordered metric structures in attitudes. *Australian Journal of Psychology* 50, 199–204.
- Myung, I.J., 2000. Importance of complexity in model selection. *Journal of Mathematical Psychology* 44, 190–204.
- Myung, I.J., Foster, M.R., Browne, M.W. (Guest Eds.), 2000. Special issue on model selection. *Journal of Mathematical Psychology*, 44.
- Nickerson, C.A., McClelland, G.H., 1988. Extended axiomatic conjoint measurement: A solution to the methodological problem in studying fertility related behaviors. *Applied Psychological Measurement* 12, 129–153.
- Park, S.K., Miller, K.W., 1988. Random number generators: Good ones are hard to find. *Communications of the ACM* 31, 1192–1201.
- Person, H.B., Barron, F.H., 1978. Polynomial psychophysics of group risk perception. *Acta Psychologica* 42, 421–428.
- Robertson, T., Wright, F.T., Dykstra, R.L., 1988. *Order Restricted Statistical Inference*. John Wiley and Sons, Chichester.
- Scheiblechner, H., 1999. Additive conjoint isotonic probabilistic models (ADISOP). *Psychometrika* 64, 295–316.
- Schwarz, W., 1990. Experimental and theoretical results for some models of random-dot-pattern discrimination. *Psychological Research* 52, 299–305.
- Scott, D., 1964. Measurement models and linear inequalities. *Journal of Mathematical Psychology* 1, 233–247.
- Sherman, B.F., 1977. An algorithm for finite conjoint additivity. *Journal of Mathematical Psychology* 16, 204–218.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P., vander Linde, A., 2001. Bayesian measures of model complexity and fit. *Journal of the Royal Statistics Society, Series B*, forthcoming.

- Tierney, L., 1994. Exploring posterior distributions with Markov chains (with discussion). *Annals of Statistics* 22, 1701–1762.
- Tversky, A., 1967. Additivity, utility, and subjective probability. *Journal of Mathematical Psychology* 4, 175–201.
- Ullrich, J.R., Painter, J.R., 1974. A conjoint measurement analysis of human judgement. *Organizational Behavior and Human Performance* 12, 50–61.
- Ullrich, J.R., Wilson, R.E., 1990. CPCJM—A set of programs for checking polynomial conjoint measurement and additivity axioms of 3-dimensional matrices. *Applied Psychological Measurement* 14, 433–434.
- Ullrich, J.R., Wilson, R.E., 1993. A note on the exact number of two- and three-way tables satisfying conjoint measurement and additivity axioms. *Journal of Mathematical Psychology* 37, 624–628.