

An iterative procedure for extracting skill maps from data

Andrea Spoto¹ · Luca Stefanutti² · Giulio Vidotto¹

Published online: 24 June 2015
© Psychonomic Society, Inc. 2015

Abstract The methodologies for the construction of a knowledge structure mainly refer to the query to experts, the skill maps, and the data-driven approaches. This last method is of growing interest in recent literature. In this paper, an iterative procedure for building a skill map from a set of data is introduced. This procedure is based on the minimization of the distance between the knowledge structure delineated by a given skill map and the data. The accuracy of the proposed method is tested through a number of simulation studies where the amount of noise in the data is manipulated as well as the kind of structure to be reconstructed. Results show that the procedure is accurate and that its performance tends to be sufficiently stable even with high error rates. The procedure is compared to two already-existing methodologies to derive knowledge structures from a set of data. The use of the corrected Akaike Information Criterion (AICc) as a stopping criterion of the iterative reconstruction procedure is tested against the *app* criterion introduced by Schrepp. Moreover, two empirical applications on clinical data are reported, and their results show the applicability of the procedure.

Keywords Skill maps · Knowledge space theory · Data-driven construction methods · Conjunctive model · Disjunctive model

Introduction

The present paper aims to introduce a new methodology within the framework of the Knowledge Space Theory (KST; Doignon & Falmagne, 1985, 1999; Falmagne & Doignon, 2011) to derive a skill map from a data set. A skill map is a function that assigns to each item the set of abilities needed to solve it. KST has been mainly applied in the assessment of knowledge of students. One of its most successful applications is the software ALEKS (Assessment and LEarning in Knowledge Spaces; Doignon & Falmagne, 1999; Falmagne, Doble, Albert, Eppstein, & Hu, 2013; Falmagne & Doignon, 2011). Nevertheless, in recent years, the KST approach has been applied in clinical contexts in order to carry out an effective and adaptive assessment of the clinical state of patients (Bottesi et al., 2014). Such applications have shown both that the KST approach could adequately fit clinical requirements, and that the possibility to further theoretically develop these innovative applications deserves a great attention.

In KST, some basic concepts are defined: a *knowledge domain* is the entire set Q of items that can be asked about a specific topic; a *knowledge state* is the subset K of the knowledge domain that a person masters; a *knowledge structure* is a pair (Q, \mathcal{K}) where \mathcal{K} is the collection of all knowledge states existing in a population, containing at least the empty set (\emptyset) and the total set (Q).

One of the core issues related to KST is the construction of knowledge structures. Generally speaking, three main areas are usually explored in addressing this task: the first

✉ Andrea Spoto
andrea.spoto@unipd.it

Luca Stefanutti
luca.stefanutti@unipd.it

Giulio Vidotto
giulio.vidotto@unipd.it

¹ Department of General Psychology, University of Padua, Via Venezia 8, 35131 Padova, Italy

² FISPPA Department, University of Padua, Via Venezia 8, 35131 Padova, Italy

way is the *query to experts* procedure (Cosyn & Thiéry, 2000; Kambouri et al., 1993; Koppen, 1993; Koppen & Doignon, 1990; Stefanutti & Koppen, 2003); the second methodology refers to the *skill maps* approach (Doignon, 1994; Düntsch & Gediga, 1995; Heller et al., 2013; Korossy, 1997; Spoto et al., 2010, 2012); the third one—of growing interest in recent literature—is the *data-driven methodology* (Schrepp, 1999b, 2002, 2003; Ünlü et al., 2013). In this article, a procedure is proposed that combines the skill map and the data-driven approaches. The skill map approach, introduced to allow KST to move beyond the mere behavioral perspective (Doignon, 1994; Doignon & Falmagne, 1999; Falmagne et al., 1990), represents an appealing methodology to build knowledge structures. It is based on the possibility of assigning to each item $q \in Q$ a subset $T \subseteq S$ of skills needed to master q . For instance, consider the following four items in elementary arithmetic:

1. $86 + 14 = ?$
2. $100 - 14 = ?$
3. $(86 + 14)/4 = ?$
4. $(100 - 14)/2 = ?$

It can be reasonably assumed that at least three skills are needed to solve one or more of the four items. More specifically, the following skills can be defined:

- a) sum two integer numbers;
- b) subtract two integer numbers;
- c) divide two integer numbers.

It can be argued that item 1 requires skill a , item 2 requires skill b , whereas item 3 requires both skills a and c , and item 4 requires both skills b and c . Starting from such assignment, it is possible to deduce which knowledge states appear in the structure (as explained in Section 2). One of the crucial aspects to be considered in performing the item-skill association is the rule, here named *model*, about the relation between items and skills. A skill map is a triple (Q, S, τ) where Q is a nonempty set of items, S is a nonempty set of skills, and τ is a mapping from the set Q of items to the power set of S , assigning to each item a nonempty set of skills. Three different models are usually taken into account: the *disjunctive*, the *conjunctive*, and the *competency model* (Doignon & Falmagne, 1999; Korossy, 1997). In the disjunctive model, in order to master q , it is sufficient to have one of the skills in $\tau(q)$. In the conjunctive model, on the contrary, all the skills assigned to an item are needed to master it, this model applies, for instance, to the example given above. For a detailed explanation of these two models as well as of the competency model, which is beyond the scope of this article, refer to, e.g., Doignon and Falmagne 1999. The data-driven methodology refers to the possibility to generate a knowledge structure starting from a set of collected data. All the data-

driven methods available at the present for the construction of knowledge structures can be subdivided into two main categories (Falmagne et al., 2013): on one hand, we have the Boolean analysis of questionnaires methods, aimed at building an implication relation among the items of a questionnaire (Sargin & Ünlü, 2009; Schrepp, 1999b); on the other hand, the second group of methods derives structures directly from data (e.g.; Desmarais & Maluf, 1996; Schrepp, 1999a). Both methodologies follow approximately the same three-step procedure: the first step consists of constructing a set of relations/knowledge structures; the second step tests such models according to a set of fitting criteria; in the last step, the best fitting model is selected as the best representation of the latent structure. Examples of the first class of methods are the Item Tree Analysis (ITA; Ünlü & Albert, 2004; vanLeeuwe, 1974) and the Inductive Item Tree Analysis (IITA; Schrepp, 2002, 2003). Examples of the second class of methods can be found, for example, in Schrepp (1999a). In the present paper, we propose a method to derive a skill map from an empirical data set. The method, which is the first attempt in literature to carry out this task, consists of first deriving a knowledge structure from data, and then deriving the skill map that corresponds to that structure. The next sections present some further details on the main concepts of KST, illustrate the proposed procedure, and, finally review a simulation study and two practical applications on real data to show how the procedure works and its pros and cons.

Extracting knowledge structures from skill maps

As already stated in the Introduction, a skill map is a triple (Q, S, τ) where Q is a nonempty set of items, S is a nonempty set of skills, and τ is a mapping from Q to $2^S \setminus \{\emptyset\}$, assigning to each item a nonempty set of skills. Considering again the example proposed in the Introduction, if we set $Q = \{1, 2, 3, 4\}$ and $S = \{a, b, c\}$, then the displayed in Table 1 can be obtained.

A skill map allows two different interpretations, corresponding to two different models. According to the conjunctive model, in order to master an item q a student must possess all the skills in $\tau(q)$. Therefore, if $T \subseteq S$ is the set

Table 1 Skill map used in the example (see the text)

Item(q)	$\tau(q)$
1	{ a }
2	{ b }
3	{ a, c }
4	{ b, c }

of all skills possessed by a student, then, the set of items mastered by this student will be:

$$K = \{q \in Q \mid \tau(q) \subseteq T\}. \quad (1)$$

This model applies to the example provided in the Introduction. Suppose that $\{a, b\}$ is the set of skills possessed by a student. By an application of Eq. 1 it is possible to conclude that the knowledge state of that student is $\{1, 2\}$. If Eq. 1 is applied to every possible subset $T \subseteq S$, a collection \mathcal{C} of knowledge states is obtained. This collection is called the knowledge structure delineated by skill map τ via the conjunctive model. In the example at hand, the delineated knowledge structure is:

$$\mathcal{C} = \{\emptyset, \{1\}, \{2\}, \{1, 2\}, \{1, 3\}, \{2, 4\}, Q\}. \quad (2)$$

It can be observed that the intersection of any two sets in \mathcal{C} is again in \mathcal{C} , i.e., the structure is closed under intersection and it is named *closure space*. This is not a coincidence, but a precise property of knowledge structures delineated via the conjunctive model. According to the disjunctive model, in order to master an item, q , a student must possess at least one of the skills in $\tau(q)$. Therefore, if $T \subseteq S$ is the set of all skills possessed by a student, then the set of items mastered by this student will be:

$$K = \{q \in Q \mid \tau(q) \cap T \neq \emptyset\}; \quad (3)$$

The collection of all knowledge states obtained in this way is the knowledge structure \mathcal{K} delineated via the disjunctive model. Knowledge structures delineated via the disjunctive model are closed under union, i.e., every union of sets in \mathcal{K} produces a set already contained in \mathcal{K} ; such structures are named *knowledge spaces*. To give an example, consider the skill map (Q, S', τ') , where $S' = \{a', b', c'\}$, and τ' is such that $\tau'(1) = \{a'\}$, $\tau'(2) = \{b'\}$, $\tau'(3) = \{a', c'\}$, and $\tau'(4) = \{b', c'\}$. Then, the knowledge structure \mathcal{K} delineated by this skill map via the disjunctive model is:

$$\mathcal{K} = \{\emptyset, \{1, 3\}, \{2, 4\}, \{3, 4\}, \{1, 3, 4\}, \{2, 3, 4\}, Q\}. \quad (4)$$

An important relationship exists between the knowledge space \mathcal{K} and the closure space \mathcal{C} delineated by the same skill map (Q, S, τ) . In fact, \mathcal{C} can be obtained from \mathcal{K} as the set of the complements of the states in \mathcal{K} , known as the dual of \mathcal{K} :

$$\mathcal{C} = \{Q \setminus K \mid K \in \mathcal{K}\}.$$

To summarize, the knowledge space \mathcal{K} delineated by a skill map τ via the disjunctive model and the closure space \mathcal{C} delineated by the same skill map via the conjunctive model are dual of one another.

Deriving skill maps from knowledge structures

The material in the previous section was based on the assumption that a skill map can be constructed starting from theoretical considerations about the items. The knowledge structure on the set of items is then derived from the skill map. In this section, we consider the opposite situation: a knowledge structure is available and a skill map delineating this structure has to be identified. This is possible if the knowledge structure is closed under either union, intersection or both. The former case is considered first.

Any knowledge space defined on a finite set Q , has a *basis*, which is the minimal set of states that cannot be obtained as unions of other states. Starting from the basis, it is possible to reconstruct the knowledge space by closing the basis under set union. In the example above, it is possible to see that the basis $\mathcal{B}_{\mathcal{K}}$ of Eq. 4 is:

$$\mathcal{B}_{\mathcal{K}} = \{\{1, 3\}, \{2, 4\}, \{3, 4\}\}.$$

Theorem 6.3.8 of Falmagne and Doignon (2011) asserts that any knowledge space is delineated by some (minimal) skill map if and only if it has a basis whose cardinality equals that of the set of skills. A strong relationship exists between the *basis* of a knowledge space and the smallest skill map delineating it.

Corollary 1 Assume that \mathcal{K} is a space with basis \mathcal{B} , and S an arbitrary set such that there is a bijective correspondence $f : S \rightarrow \mathcal{B}$. Then the space delineated by the disjunctive skill map $\tau(q) = \{s \in S : q \in f(s)\}$ is \mathcal{K} .

Proof First, it can be seen that $\tau(q) = \{s \in S : q \in f(s)\} \Leftrightarrow s \in \tau(q) \Leftrightarrow q \in f(s)$. Then, by the disjunctive model, we know that $K = \{q \in Q : \tau(q) \cap T \neq \emptyset\}$, i.e., Eq. 3. By applying Eq. 3 to the singleton $\{s\}$ we have $\{q \in Q : \tau(q) \cap \{s\} \neq \emptyset\} = \{q \in Q : s \in \tau(q)\} = \{q \in Q : q \in f(s)\}$. By applying the same procedure to all the singletons, all the element of the basis \mathcal{B} are reproduced. Then, \mathcal{K} is obtained by closure under union of \mathcal{B} . \square

To give an example of the application of Corollary 1, let $S = \{a, b, c\}$, and define the bijection $f : S \rightarrow \mathcal{B}_{\mathcal{K}}$ such that $f(a) = \{1, 3\}$, $f(b) = \{2, 4\}$, and $f(c) = \{3, 4\}$. Now, setting $\tau(q) = \{s \in S : q \in f(s)\}$ for each item $q \in Q$ the skill map in Table 1 is obtained. We recall that this is exactly the skill map delineating \mathcal{K} under the disjunctive model.

Suppose now that a knowledge structure \mathcal{C} is closed under set intersection. The dual of \mathcal{C} is a knowledge space and by Corollary 1 it is delineated by a minimal skill map τ via the disjunctive model. This implies that \mathcal{C} is delineated by τ via the conjunctive model. Therefore, if in a concrete situation the objective is to find the skill map delineating \mathcal{C}

via the conjunctive model, the following steps provide the desired results:

1. the dual \mathcal{K} of \mathcal{C} is computed;
2. the basis $\mathcal{B}_{\mathcal{K}}$ of \mathcal{K} is obtained;
3. a set S of abstract skills is defined such that there is a bijection $f : S \rightarrow \mathcal{B}$;
4. the skill map τ is defined such that $\tau(q) = \{s \in S : q \in f(s)\}$ for each $q \in Q$.

For the reasons explained above, the skill map τ delineates the structure \mathcal{C} via the conjunctive model. To give an example, suppose that the knowledge structure is as in Eq. 2. Since \mathcal{C} is closed under intersection, the procedure described above can be applied for deriving the corresponding skill map. To do so, it is necessary to compute first the dual of \mathcal{C} , which is already defined in Eq. 4 and its basis $\mathcal{B}_{\mathcal{K}}$. Therefore, since $\mathcal{B}_{\mathcal{K}}$ contains three states, the set S of skills will also contain three skills that we name a , b , and c . Now, by applying step 4 of the procedure, the skill map τ displayed in Table 1 is obtained. Of course the interpretation of the skill map is in this case conjunctive.

The basic local independence model

So far, we introduced the main deterministic concepts of KST. In practical applications, however, probabilistic models are needed. One of the most common probabilistic models applied to knowledge structures is the basic local independence model (BLIM; Falmagne & Doignon, 1988a; Doignon & Falmagne, 1999; Falmagne & Doignon, 2011). BLIM parameters are the η_q (i.e., a lucky guess), the β_q (i.e., a careless error), and π_K (i.e., the probability assigned to each state of the structure). In the BLIM, the answers to the items are locally independent given the particular knowledge state of a person in the sense displayed by Eq. 5. The response pattern of an individual is represented by the set $R \subseteq Q$ of all items that obtained a correct response. This model has been applied in several contexts (e.g., Falmagne et al., 1990; Stefanutti, 2006; Spoto et al., 2010). Starting from the probabilistic structure (Q, \mathcal{K}, π) the probability of a response pattern $R \subseteq Q$ is (Falmagne & Doignon, 1988b):

$$p(R) = \sum_{K \in \mathcal{K}} \rho(R, K) \pi(K),$$

where ρ is a function assigning to each R its conditional probability given a state K . Notice that the knowledge states are latent, so they are not directly observable; what is observed is the response pattern R , which may be the result of some lucky guesses and/or careless errors on single items. The function ρ satisfies local independence for the items in

Q , thus $\rho(R, K)$ is determined on the basis of the two error parameters of the items, β_q , and η_q by Eq. 5:

$$\rho(R, K) = \left[\prod_{q \in K \setminus R} \beta_q \right] \left[\prod_{q \in K \cap R} (1 - \beta_q) \right] \times \left[\prod_{q \in R \setminus K} \eta_q \right] \left[\prod_{q \in \overline{K \cup R}} (1 - \eta_q) \right]. \quad (5)$$

The introduction of these notions about KST allows considering in more detail some of the already-existing procedures to build a knowledge structure from a set of data. More specifically, in the next subsection, the critical elements of both Boolean analysis of questionnaires and IITA are explored in order to clarify the rationale for looking for a new method able to cope with all these limitations.

Limitations of existing data-driven methods

In general, from a cognitive perspective, the most important limitation of both query and data-driven procedures is that they work at the items level, leaving out any information concerning the skills level. However, what people learn is skills, not items. Items (regarded as questions or problems) are means for ascertaining whether skills are mastered or not by individuals. Shifting the attention from the behavioral level of the items to the competence level of the skills offers a possible cognitive interpretation of the response behavior of an individual. Referring to the example proposed in the Introduction, both query and data-driven procedures would reconstruct the structure, but do not formally/explicitly provide information about the cognitive aspects that make item 1 less demanding than item 3. Importantly, a great amount of research concerning the role of the skills or abilities to master items has been conducted in the area of cognitive diagnosis models (CDM; e.g., DeLaTorre & Douglas, 2008; Templin & Henson, 2006), and it has also been deeply investigated in KST (see e.g., Doignon, 1994; Düntsch & Gediga, 1995; Heller et al., 2013; Korossy, 1997). Furthermore, considering the main data-driven approaches, some other considerations have to be taken into account. In his paper, Schrepp (1999a) indicates at least two main critical aspects of his procedure based on sorting the observed response patterns by decreasing frequency and then adding to the reconstructed structure one pattern at a time. At each step, a fit index is computed and the model that displays the minimum value of such index is selected as the correct knowledge structure. The two main criticisms of the procedure introduced by Schrepp are: the number of observed patterns necessary to reliably reconstruct the structure increases heavily with the cardinality of Q , thus, a researcher who wants to use such a procedure may have

to face the impossibility of collecting such immense data sets; if the states of the underlying knowledge structure do not have equal probability, the accuracy of the procedure is influenced negatively. Furthermore, the procedure does not account for any state that is not observed in terms of response pattern, and this could be a problem. In fact, suppose a population is characterized by a knowledge structure \mathcal{K} on the set Q of items, and a sample of response patterns $R \subseteq Q$ is available. Some of the states $K \in \mathcal{K}$ might be not contained in the sample for at least two different reasons: (1) π_K is small, and for this reason K does not belong to the sample at hand, or (2), due to lucky guesses and careless errors, none of the response patterns in the sample equals K . If either of these two cases occurs, Schrepp's procedure fails to reconstruct the "true" structure \mathcal{K} . In other words, if the response pattern corresponding to a latent state of the knowledge structure is not observed in the sample, it will never be included in the structure. Moreover, given that the introduction of a pattern into the structure follows the frequency criterion, the procedure does not guarantee to produce a knowledge structure (since either the total set Q or the empty set \emptyset may have a very low frequency in the sample at hand). Another important criticism of this procedure is that it both underestimates the error rates of the items and assumes that such values are equal for all the items of the domain (Schrepp, 1999a).

The IITA procedure, on the other hand, is conceived to build knowledge structures that are closed under both set union and set intersection, that is, it systematically produces *quasi-ordinal knowledge spaces* (Falmagne & Doignon, 2011). In this respect, one main criticism arises: the assumption of having a structure both \cap - and \cup -closed may be too strong in several applicative contexts. Given the limitations of the existing procedures, we aimed at proposing a new approach able to cope with most of the critical issues described above. More specifically, we aimed at obtaining a structure that could be traced back to a skill map (i.e., it has to be closed under intersection, union, or both) and does not present IITA strong assumptions. In the next section, the procedure used to derive the knowledge structure and the skill map is described.

Procedure

The procedure needed to exploit the task of building a skill map out of a set of data is supposed to address the following two problems:

- i) to build a chain of nested closure spaces on the basis of a given criterion (as described below);
- ii) to select one of the closure spaces according to an appropriate criterion.

The procedure would then end up with the best fitting closure space from which, following the principles described in the previous sections, a conjunctive skill map is obtained. Let $\mathcal{R} = 2^Q$ be the power set on Q . The data set consists of N response patterns $R \in \mathcal{R}$. We assume that for every theoretically observable response pattern R , the observed frequency $F(R)$ is known (with $\sum_{R \in \mathcal{R}} F(R) = N$).

To address problem i), a distance is minimized between the structure \mathcal{K} and the data \mathcal{R} . Suppose that a student in knowledge state $K \in \mathcal{K}$ exhibits response pattern $R \in \mathcal{R}$. Then, the distance between R and K is at least $d_{\min}(R, \mathcal{K}) = \min\{|R \Delta K| : K \in \mathcal{K}\}$, with

$$\Delta = (RK) \cup (KR)$$

denoting the symmetric distance between R and K . The average minimum distance for the whole data set \mathcal{R} is:

$$d(\mathcal{R}, \mathcal{K}) = \frac{1}{N} \sum_{R \in \mathcal{R}} F(R) d_{\min}(R, \mathcal{K})$$

The procedure is aimed at finding the structure \mathcal{K} for which $d(\mathcal{R}, \mathcal{K})$ is the smallest possible. The starting structure is the simplest possible closure space, i.e., $\mathcal{C}_0 = \{\emptyset, Q\}$. Then, a chain of closure spaces with a strictly increasing cardinality is built. In every single step $n > 0$ of the iterative procedure, the transition from \mathcal{C}_{n-1} to \mathcal{C}_n is carried out by calculating, for each observed response pattern $R \in \mathcal{R} \setminus \mathcal{C}_{n-1}$, the distance $d(\mathcal{R}, \mathcal{D}_R)$ where \mathcal{D}_R is the closure under intersection of $\mathcal{C}_{n-1} \cup \{R\}$. This computation is repeated for every $R \in \mathcal{R} \setminus \mathcal{C}_{n-1}$. The pattern R_{best} , for which $d(\mathcal{R}, \mathcal{D}_R)$ is the minimum, is considered. At this point, \mathcal{C}_n is set equal to $\mathcal{D}_{R_{best}}$. The procedure ends at step m such that $\mathcal{R} \setminus \mathcal{C}_{m-1} = \emptyset$, i.e., when there are no patterns left to add. The procedure, thus, minimizes at each step the average minimum distance between the built structure and the observed data, respecting the condition of closure under intersection. Furthermore, it must be stressed that at each step of the procedure, this distance monotonically decreases and attains the value 0 when all the patterns are included in the model; this turns out to be nothing but the closure under intersection of the set of response patterns. For this reason, the distance cannot be considered a good model selection criterion, as it would always select the most complex structure. The establishment of a reliable and adequate criterion for selecting the best model both in terms of fit and parsimony is then a crucial issue and represents the second task of the procedure.

Problem ii) is mainly a task of model selection. In the literature, two of the main indexes used to evaluate the goodness of a model are the Akaike Information Criterion (AIC), and the corrected AIC (AICc; Brockwell & Davis, 1991; Burnham & Anderson, 2002) which can be used conveniently with small sample sizes. Nevertheless, the application of these criteria to KST models may be problematic because, even with item sets of moderate size,

the sample will usually be small compared to the total number of observables. For instance, with $|Q| = 20$, the number of observables is 2^{20} , which is more than one million different response patterns. Every reasonable sample size is small compared to this number. Schrepp (1999a) introduced the *app* stopping criterion which is defined as follows:

$$app(\mathcal{K}_L, \mathcal{R}) = \frac{\sum_{R \in \mathcal{R}} (F(R) - F_L(R))^2}{|\mathcal{R}|}$$

Such an index refers to the distance between the reconstructed knowledge structure \mathcal{K}_L , at each step of the procedure, and the set of observed response patterns \mathcal{R} . This distance is the ratio between the squared difference of relative frequencies $F(R)$ of the patterns and theoretical frequency $F_L(R)$ expected assuming \mathcal{K}_L to be the correct structure, and the cardinality of \mathcal{R} . The smaller the value of *app*, the better the approximation of the model to the data is. The step in which *app* reaches its minimum corresponds to the best fitting model. In our application, we selected this index as a stopping criterion. Its computation requires that the theoretical frequencies $F_L(R)$ are available. These expectations can be obtained by estimating the BLIM for the knowledge structure \mathcal{K}_L .

Once the structure is selected, since it is a closure space, the basis of its dual can be obtained, and, from this basis, the corresponding conjunctive skill map can be derived. Like Schrepp's procedure, the proposed one is carried out in a number of steps. In each step i a knowledge structure is specified, which is a strict superset of the one obtained in step $i - 1$. A stopping criterion is used for deciding when the procedure should terminate. The proposed procedure differs from the Schrepp's one with respect to three basic aspects: (1) the next response patterns to be added to the structure are selected by a minimum distance, rather than by the maximum frequency criterion; (2) the structure is closed under intersection in every single step; in the proposed procedure the *app* criterion is based on the careless error and lucky guess parameter estimates obtained by fitting the BLIM to the data, meaning that β and η parameters can be different for different items.

So far, the procedure was described for extracting a conjunctive skill map from data. Nevertheless, if the aim is to extract a disjunctive skill map, this procedure can be adapted with minimal effort. In fact, it is sufficient to replace closure under intersection by closure under union at each step n of the procedure. The remaining parts of the procedure are exactly the same. Of course, there is the possibility of performing both closures at each step, thus obtaining, as the final result, a quasi-ordinal knowledge space as in the IITA procedure.

Simulation study

In this section, we introduce a simulation study carried out to test the efficiency and accuracy of the iterative procedure to build a closure space starting from a set of observed data. The performance of the Data Driven Skill Map Extraction Procedure (D-SMEP) is compared to that of two of the main alternative procedures, i.e., the Schrepp's method and the Inductive Item Tree Analysis (IITA).

Method

For the simulation, the following procedure was applied. We aimed at testing the performance of the three procedures by varying both the characteristics of the given "true" structure, and the magnitude of the careless error and lucky guess probabilities of the items. As it has been described in Section 2, Schrepp's procedure extracts a knowledge structure that needs being closed under neither union nor intersection; on the other hand, IITA extracts quasi-ordinal knowledge spaces, and finally, the procedure proposed in this article extracts a closure space (respectively, knowledge space) and the corresponding conjunctive (respectively, disjunctive) skill map. For this reason, the three procedures were compared to one another with respect to their capability of reconstructing (1) a "true" structure neither closed to union nor intersection, (2) a "true" structure closed under both union and intersection, and, finally, (3) a "true" structure closed under intersection. The error rates of the items were randomly selected within an interval ranging between 0 and a maximum chosen among the three values 0.1, 0.2, and 0.3. The probability assigned to each state of the structure (i.e., the parameter π_K) was a random number in the open interval (0, 1). This scheme leads to non-uniform sampling from the distributions over the knowledge states (Smith & Tromble, 2004); nevertheless, given that these values are assigned only once, this distortion in the simulation at hand is negligible. The π_K parameters were then normalized to sum up to 1. A total of 3 types of knowledge structures \times 3 maximum error rates were considered. For each condition, the following procedure was carried out:

- i) A number of 100 random samples each containing 1000 response patterns were generated by parametric bootstrap (Efron & Tibshirani, 1993), where the probabilistic model was the BLIM. In all replications, the same fixed knowledge structures \mathcal{K}_t on eight items containing 50 states for the conditions (1) and (2) and containing 48 states for the condition (3), were used as the true structure.
- ii) The three iterative procedures were applied to every single data set, obtaining a structure for each of the 100 data sets in each of the nine conditions.

iii) With respect to the new reconstruction procedure, at every step both the corrected AIC (Brockwell & Davis, 1991; Burnham & Anderson, 2002), the *app* index (Schrepp, 1999a) and a Distance from the True Structure (DTS) were computed. The DTS was obtained through the following:

$$\begin{aligned} \text{DTS}(\mathcal{K}_t, \mathcal{C}_n) &= d(\mathcal{K}_t, \mathcal{C}_n) + d(\mathcal{C}_n, \mathcal{K}_t) \\ &= \frac{1}{|\mathcal{K}_t|} \sum_{K \in \mathcal{K}_t} d_{\min}(K, \mathcal{C}_n) \\ &\quad + \frac{1}{|\mathcal{C}_n|} \sum_{C \in \mathcal{C}_n} d_{\min}(C, \mathcal{K}_t) \end{aligned}$$

where \mathcal{C}_n is the structure obtained at step n of the procedure. The DTS attains the value 0 if and only if $\mathcal{K}_t = \mathcal{C}_n$. On the other hand, given that by definition both the empty set (\emptyset) and the total set Q belong to both \mathcal{K}_t and \mathcal{C}_n , the maximum possible value of DTS is equal to $|Q|/2$. In fact, the maximum value of the minimum distance between a given state in \mathcal{K}_t and a state in \mathcal{C}_n is half the number of items in Q . Moreover, if the DTS between the two structures \mathcal{K}_t and \mathcal{C}_n is d , then on the average any given state $K \in \mathcal{K}_t$ differs from its closest state $C_i \in \mathcal{C}_n$ by d items.

With respect to the other two methodologies (IITA and Schrepp), for each condition, all the steps of the procedures were carried out in order to obtain a structure according to the assumptions of each method.

- iv) The following additional statistics were computed for each of the three procedures, in each of the nine conditions:
- The average proportion of states of the “true” structure contained in the reconstructed structure (true positive states).
 - The average proportion of states not contained in the “true” structure, but contained in the reconstructed structure (false positive states).
 - The average distance $d(\mathcal{K}_t, \mathcal{C}_n)$ between the “true” structure and the reconstructed one.
 - The average distance $d(\mathcal{C}_n, \mathcal{K}_t)$ between the reconstructed structure and the “true” one. It should be observed that, in general $d(\mathcal{C}_n, \mathcal{K}_t) \neq d(\mathcal{K}_t, \mathcal{C}_n)$ (see e.g., Falmagne & Doignon, 2011).
 - The average DTS.

Concerning Schrepp’s procedure, all these indexes were computed on the structures obtained at the iteration in which the minimum *app* was obtained; the IITA procedure does not allow to compute the *app* index. Therefore, for this procedure, these statistics were computed on the obtained model; in D-SMEP these statistics were computed

on the structures obtained when the *app* reached its own minimum. Additionally, for the proposed procedure, the statistics were computed also on the structure corresponding to the minimum of the AICc.

Results

Tables 2, 3 and 4 display the main results of the simulations carried out to compare the efficiency and accuracy of the three procedures.

Table 2 displays the results obtained in reconstructing a structure neither closed to union nor to intersection at increasing error rates. In these conditions, the best performing procedure is the one proposed by Schrepp. In fact, it maximizes the proportion of true states, minimizes the amount of false states in the reconstructed structure as well as the distances between the true and the reconstructed models. In this set of conditions, the underlying knowledge structure cannot be obtained through any skill map (since no closure is present). Nevertheless, it has to be stressed that the proposed method (using the *app* index as a stopping criterion) performs almost as well as the Schrepp’s one. Thus, it produces a structure that is adequately close to the true one, and it can be conveniently described by a skill map.

Table 2 The performances of the three procedures in reconstructing a structure neither closed under union nor intersection. The table displays the proportion of true states included in the reconstructed structure (True Pos), the proportion of the false states in the reconstructed structure (False Pos), the average distance between the true structure and the reconstructed structure ($d(\mathcal{K}_t, \mathcal{C}_n)$), the average distance between the reconstructed structure and the true structure ($d(\mathcal{C}_n, \mathcal{K}_t)$), and the DTS (notice that each of the three distances take on values in the interval $[0, |Q|/2]$). For our procedure, we reported these statistics with respect to two different model selection criteria: the AICc (D-SMEP AICc) and the *app* index (D-SMEP *app*)

Error rate	Method	True Pos	False Pos	$d(\mathcal{K}_t, \mathcal{C}_n)$	$d(\mathcal{C}_n, \mathcal{K}_t)$	DTS
.10	Schrepp	0.997	0.168	0.003	0.168	0.171
	IITA	0.320	0.693	0.879	0.697	1.576
	D-SMEP AICc	0.979	0.788	0.021	0.935	0.956
	D-SMEP <i>app</i>	0.980	0.786	0.020	0.932	0.953
.20	Schrepp	0.990	0.483	0.010	0.491	0.501
	IITA	0.389	0.693	0.775	0.723	1.498
	D-SMEP AICc	0.789	0.752	0.216	0.871	1.087
	D-SMEP <i>app</i>	0.985	0.789	0.015	0.939	0.954
.30	Schrepp	0.952	0.680	0.048	0.732	0.779
	IITA	0.564	0.711	0.559	0.777	1.336
	D-SMEP AICc	0.274	0.602	1.004	0.632	1.636
	D-SMEP <i>app</i>	0.999	0.804	0.001	0.963	0.964

Table 3 The performances of the three procedures in reconstructing a structure closed under both union and intersection (notice that $d(\mathcal{K}_t, \mathcal{C}_n)$, $d(\mathcal{C}_n, \mathcal{K}_t)$ and DTS take on values in the interval $[0, |Q|/2]$)

Error Rate	Method	True Pos	False Pos	$d(\mathcal{K}_t, \mathcal{C}_n)$	$d(\mathcal{C}_n, \mathcal{K}_t)$	DTS
.10	Schrepp	0.991	0.004	0.009	0.004	0.012
	IITA	1	0.001	0	0.001	0.001
	D-SMEP AICc	0.393	0.184	0.699	0.184	0.883
	D-SMEP <i>app</i>	0.976	0.439	0.024	0.457	0.481
.20	Schrepp	0.987	0.082	0.013	0.082	0.094
	IITA	0.935	0.046	0.065	0.046	0.111
	D-SMEP AICc	0.482	0.156	0.542	0.161	0.704
	D-SMEP <i>app</i>	0.977	0.414	0.024	0.462	0.485
.30	Schrepp	0.982	0.424	0.012	0.440	0.459
	IITA	0.992	0.376	0.001	0.428	0.428
	D-SMEP AICc	0.421	0.138	0.735	0.139	0.873
	D-SMEP <i>app</i>	0.991	0.566	0.009	0.642	0.651

Our procedure provides adequate results in all the three error rate conditions when the *app* is used as a stopping criterion, while the use of AICc seems to be much more critical when the error tends to increase. Concerning Tables 3 and 4, analogous conclusions can be drawn. For this reasons, it was decided to use the *app* as a stopping criterion. With respect to the performance of the IITA procedure, it has to be stressed that its performance is systematically lower

Table 4 The performances of the three procedures in reconstructing a structure closed under intersection (notice that $d(\mathcal{K}_t, \mathcal{C}_n)$, $d(\mathcal{C}_n, \mathcal{K}_t)$ and DTS take on values in the interval $[0, |Q|/2]$)

Error Rate	Method	True Pos	False Pos	$d(\mathcal{K}_t, \mathcal{C}_n)$	$d(\mathcal{C}_n, \mathcal{K}_t)$	DTS
.10	Schrepp	0.994	0.016	0.006	0.016	0.022
	IITA	0.588	0.492	0.492	0.568	1.060
	D-SMEP AICc	0.981	0.405	0.019	0.431	0.450
	D-SMEP <i>app</i>	0.965	0.361	0.035	0.383	0.418
.20	Schrepp	0.991	0.408	0.009	0.415	0.424
	IITA	0.459	0.636	0.634	0.805	1.439
	D-SMEP AICc	0.798	0.343	0.209	0.369	0.575
	D-SMEP <i>app</i>	0.969	0.599	0.031	0.699	0.730
.30	Schrepp	0.969	0.604	0.033	0.677	0.709
	IITA	0.604	0.584	0.525	0.676	1.201
	D-SMEP AICc	0.483	0.273	0.627	0.316	0.944
	D-SMEP <i>app</i>	0.966	0.728	0.034	0.926	0.960

when the true model is not a quasi-ordinal knowledge space. On the other hand, it is the best performing method when the underlying true structure satisfies this condition. It is important to stress that both Schrepp's method and D-SMEP do provide adequate results in a quite stable way across different true model and error rates conditions. This consideration seems to be further supported by the low value (i.e., always less than 1 item) of the DTS even with high error rates.

Discussion

In general, the results of the simulations support the goodness of the proposed reconstruction procedure, since the proportion of true and false states in the reconstructed structure as well as the average DTS are adequate in all the simulated conditions. The procedure has a better performance when the *app* index is used as a stopping criterion. Furthermore, as it could have been easily predicted, the procedure tends to perform better when careless error and lucky guess probabilities are relatively small. However, this difference in performance is not dramatic, allowing to trustfully apply the method even with high error rate.

The results suggest that the AICc criterion works adequately in some situations (namely those with low error rates), but fails in others, pointing that the *app* can be a better performing stopping criterion. Schrepp's procedure does perform very well in all conditions. Finally, the IITA procedure seems to work very well when the underlying model satisfies both closures, while it under-performs in all other cases.

With respect to D-SMEP, some criticisms have emerged around algorithm efficiency. For instance, the time required to carry out a single replication on rather small structures was about 500 seconds on a standard machine. This time is expected to increase very rapidly with the number of items. This indicates the need to improve the optimization of the computational steps.

Finally, in comparing D-SMEP with Schrepp's procedure, it could be useful to consider that they differ by three factors: (1) the selection method (minimum distance vs. maximum frequency); (2) closure under intersection vs. no closure; (3) parameter estimation method used for the computation of the *app* index (iterative vs. noniterative). This distinction would lead to a 2 (selection methods) \times 2 (closure vs. no closure) \times 2 (iterative *app* vs. non iterative *app*) factorial design. This would lead to the assessment of the other six procedures, all different from both D-SMEP and the Schrepp's one. Taking into account that differences in performance between D-SMEP and the Schrepp's procedure are fairly small in all simulations, it is not sure that this factorial study would help understanding such small differences.

In the next section, two empirical applications on clinical data are described in order to provide some sort of a tutorial about how to apply D-SMEP.

Empirical applications

To test how the reconstruction procedure works on real data sets and how it performs compared to IITA and Schrepp's procedures, responses were used from a sample of 4412 individuals (Spoto et al., 2010) to the subscales "Cleaning" and "Checking" of the Maudsley Obsessive-Compulsive Questionnaire-R (MOCQ-R; Sanavio & Vidotto, 1985). Both subscales contain eight dichotomous items investigating respectively cleaning rituals and repeated control behaviors in patients. The eight items of each scale are displayed in Table 5. The reconstruction of a structure out of these two subscales of the questionnaire might be useful to both identifying implications among the items composing each scale, and implementing an adaptive assessment for the disorder at hand.

These empirical applications have been selected since KST has recently been applied to clinical assessment through the Formal Psychological Assessment (FPA; Bottesi et al., 2014; Spoto, Bottesi, Sanavio, & Vidotto, 2013). In FPA, the main concepts of KST are formulated according to clinical psychology-specific issues: items are clinical questions included in surveys, skills are named *attributes* and refer to specific symptoms investigated by each item, careless error and lucky guess are, respectively, *false-negative* and *false-positive* rates. What remains unchanged is the methodology to obtain a structure starting from an attribute assignment through the conjunctive or disjunctive model and the probabilistic model related to the structures. One of the crucial issues in FPA is the conceptualization of the skills as the attributes, i.e., the clinical

symptoms, diagnostic criteria, or characteristics that an individual has to present in order to affirmatively answer to a clinical item. In this perspective, the response pattern of a patient is directly related to the set of clinical symptoms he/she presents.

Methods

The questionnaires were administered to participants after they signed the informed consent form. No time limit was imposed, and participants were asked to answer all items accurately. After the administration of the tests, data were collected, and the frequencies of the response patterns were computed. The data consisted of a set of response patterns $R \subseteq Q$ (where Q is the collection of items), with each response pattern with its own observed frequency $F(R)$. The three procedures for structure construction were then applied. In the proposed procedure, for each iteration, the *app* index was computed, and all the BLIM's parameters were estimated. The model with the minimum value of the *app* index was then selected and, by applying the procedure described in Section 2, the basis and the corresponding skill map were obtained. With respect to IITA and Schrepp's procedures, they were applied using the same steps described for the simulation study. Three different structures were obtained and each of them was fitted to data by an application of the BLIM. The three fitted models were then compared through their Chi-square and AIC values.

Results

An overview of the main indexes used to compare the performances of the three reconstruction procedures with the data at hand is displayed in Table 6 for both the subscales.

With respect to the "Checking" sub-scale, for the proposed procedure, the selected structure had 56 states. The

Table 5 The eight items of the "Cleaning" and "Checking" sub-scales of the MOCQ-R

Item	Cleaning	Checking
i_1	If I touch an animal, I feel contaminated	I frequently have to check things (e.g., gas or water taps, doors, etc.)
i_2	I use too much soap to wash myself	In my job I am often late because I spend time to do the same thing over and over again
i_3	I am not excessively concerned about cleanliness	I have to check a letter over and over again before post it
i_4	I am not unduly concerned about germs and diseases	One of my major problems is that I pay too much attention to detail
i_5	My hands feel dirty after touching money	My major problem is repeated checking
i_6	I take rather a long time to complete my washing in the morning	I tend to check things more than once
i_7	I use a great deal of detergents and antiseptics	I stick to a very strict routine when doing ordinary things
i_8	Hanging and folding my clothes at night takes up a long time	I spend a lot of time every day checking things over and over again

Table 6 The performances of the three procedures in reconstructing the knowledge structure for the “Checking” and “Cleaning” subscales

Checking					
	N states	χ^2	<i>df</i>	Bootstrap <i>p</i>	AIC
D-SMEP	56	518.35	184	0.0008	35,675
IITA	7	1098.50	233	< 0.0001	36,049
Schrepp	3	1368.70	237	< 0.0001	36,246
Cleaning					
D-SMEP	34	359.48	206	0.7697	36,208
IITA	39	301.54	201	0.9851	36,164
Schrepp	4	497.72	236	0.0045	36,276

Chi-square statistics indicated an unacceptable fit of the BLIM ($\chi^2(184) = 518.35$; bootstrap $p = .0008$; AIC= 35,675). Starting from the reconstructed structure (i.e., a closure space), through the steps described in Section 2, it was possible to obtain the dual of the closure space, that is, a knowledge space. Thus, the following basis was obtained:

$$\mathcal{B}_{CH} = \{\{i_2\}, \{i_5\}, \{i_7\}, \{i_2, i_8\}, \{i_2, i_3, i_8\}, \{i_1, i_2, i_8\}, \\ \{i_2, i_4, i_8\}, \{i_2, i_5, i_6, i_8\}\}$$

Finally, the conjunctive skill map displayed in Table 7 was produced.

The IITA procedure reconstructed a structure containing only seven states and the model displayed a worse fit of the BLIM ($\chi^2(233) = 1098.50$; bootstrap $p = < .0001$; AIC= 36,049). Similar results were obtained by applying Schrepp’s procedure. The reconstructed model counted only three states and, obviously, displayed the worst fit of the BLIM ($\chi^2(237) = 1368.70$; bootstrap $p = < .0001$; AIC= 36,246). It has to be stressed that the initial application of this last procedure did not reconstruct a knowledge structure, but it selected only two states excluding the total set.

Table 7 Binary representation of the skill map obtained for the empirical application with respect to the “Checking” sub-scale

Items	Attributes							
	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>	<i>h</i>
<i>i</i> ₁	0	0	0	0	0	1	0	0
<i>i</i> ₂	0	1	0	1	1	1	1	1
<i>i</i> ₃	0	0	0	0	1	0	0	0
<i>i</i> ₄	0	0	0	0	0	0	1	0
<i>i</i> ₅	0	0	1	0	0	0	0	1
<i>i</i> ₆	0	0	0	0	0	0	0	1
<i>i</i> ₇	1	0	0	0	0	0	0	0
<i>i</i> ₈	0	0	0	1	1	1	1	1

A 1 in the matrix indicates that to affirmatively answer the item in the corresponding row, the attribute in the corresponding column is needed

Such a state had to be manually added after the completion of the reconstruction procedure in order to obtain a knowledge structure. Table 8 displays the error rate estimates for the three reconstructed models for both the “Checking” and the “Cleaning” subscales. Some of the parameters are not identifiable (in the table *n.i.*) due to the particular form of the obtained structures (Spoto et al., 2012). Moreover, it can be observed that some of the probabilities are rather high, but this is in line with their interpretation in terms of false-positive and false-negative rates in a clinical questionnaire (see “Discussion”).

With respect to the “Cleaning” subscale, in running the reconstruction procedure with the data at hand, the proposed procedure selected a structure with 34 states. The Chi-square statistics indicated a good fit of the BLIM ($\chi^2(206) = 359.48$; bootstrap $p = .77$; AIC= 36,208). Starting from the reconstructed closure space, the following basis was obtained:

$$\mathcal{B}_{CL} = \{\{i_1, i_2, i_3\}, \{i_1, i_2, i_7\}, \{i_1, i_2, i_7, i_8\}, \{i_1, i_2, i_6, i_7\}, \\ \{i_1, i_2, i_4, i_7\}, \{i_1, i_2, i_5, i_7\}\}$$

Finally, the conjunctive skill map displayed in Table 9 was derived.

From the table it emerges that items *i*₁ and *i*₂ have the same set of attributes, thus they are equally informative, i.e., they are systematically present in the same states of the structure. Table 8 displays the error rate estimates. As in the previous application, some of the parameters are not identifiable. In this application, the best performing procedure happened to be the IITA that reconstructed a structure containing 39 states with a better fit of the BLIM ($\chi^2(201) = 301.54$; bootstrap $p = .99$; AIC= 36,164). Notice that even in this case, D-SMEP did not greatly differed from the best performing one. Once again, the Schrepp’s procedure reconstructed a model containing only four states and with an inadequate fit to the BLIM ($\chi^2(236) = 497.72$; bootstrap $p = .005$; AIC= 36,276).

Discussion

The results of the empirical applications of the procedure show both the applicability of the methodology and its good performance compared to the already-existing methodologies to reconstruct a knowledge structure starting from a set of data. The applications provide also some information about the possibility of interpreting the derived skill map in clinical context. More specifically, the application to the “Checking” sub-scale presents some interesting aspects that are noteworthy. First, the obtained structure counts only 56 out of the 256 potential response patterns (i.e., the power-set on the set of eight items). Furthermore, a remark has to be referred to the fact that the reconstructed skill map has as many skills as the number of items, although this does

Table 8 Estimated BLIM’s parameters β and η for each item of the “Checking” and “Cleaning” subscales of the MOCQ-R (*n.i.* means that the parameter was not identifiable)

Structures obtained by		i_1	i_2	i_3	i_4	i_5	i_6	i_7	i_8
Checking									
η	D-SMEP	<i>n.i.</i>	0.057	<i>n.i.</i>	<i>n.i.</i>	0.027	<i>n.i.</i>	<i>n.i.</i>	0.013
	IITA	0.183	0.052	<i>n.i.</i>	<i>n.i.</i>	<i>n.i.</i>	<i>n.i.</i>	0.053	<i>n.i.</i>
	Schrepp	0.187	0.051	<i>n.i.</i>	0.284	0.051	0.354	0.277	0.006
β	D-SMEP	0.258	<i>n.i.</i>	0.346	0.184	<i>n.i.</i>	0.050	<i>n.i.</i>	0.381
	IITA	<i>n.i.</i>	0.638	0.356	0.226	0.138	0.044	<i>n.i.</i>	0.398
	Schrepp	0.304	0.757	0.465	0.219	0.141	0.036	0.198	0.630
Cleaning									
η	D-SMEP	0.065	0.025	0.010	<i>n.i.</i>	<i>n.i.</i>	<i>n.i.</i>	0.009	<i>n.i.</i>
	IITA	0.037	0.028	0.139	0	0.081	0	<i>n.i.</i>	0.064
	Schrepp	0.054	0.020	0.055	0	0.258	0.132	<i>n.i.</i>	0.347
β	D-SMEP	0.719	0.525	0.233	0.449	0.335	0.315	0.568	0.331
	IITA	0.514	0.518	0.250	0.409	0.316	0.284	0.191	0.345
	Schrepp	0.764	0.666	0.254	0.453	0.324	0.349	0.689	0.312

The parameters are displayed separately for each reconstructed knowledge structure to which the BLIM was applied

not mean that items are independent. For instance, if items i_5 and i_6 are considered, it can be argued that if no error is taken into account, an affirmative response to the former implies an affirmative answer to the latter since the set of attributes referred to i_6 is strictly included in the set of attributes of i_5 .

The application to the “Cleaning” subscale, on the other hand, depicts a case in which the number of attributes is less than the number of items. Moreover, the proposed procedure identifies two items that need the same set of attributes. The critical aspect about this application is referred to the high values of the error rate estimates obtained with all three

reconstruction procedures. Response errors might have a quite different interpretation in a clinical questionnaire. They may reflect the tendency to simulate or dissimulate a symptom, depending on social desirability or other similar issues. This can lead to substantially greater error probabilities. It is noteworthy how Schrepp’s procedure fails even in this application, while the IITA reconstructs a model that better fits the BLIM. These results seem to be in contradiction with those obtained in the simulation study. A possible issue could be the characteristics of the simulated models. To verify this, 100 data sets were simulated by the BLIM, using as “true states” the knowledge structure reconstructed by IITA and as “true parameters” those estimated by fitting the BLIM to data with the knowledge structure obtained by IITA. Schrepp’s procedure was then applied to each of the 100 data sets and the DTS, TPR, and FPR statistics were computed. The mean DTS was 1.11, the mean TPR was .079, the mean FPR was 0, the mean number of states in the reconstructed structure was 3.07. These results (especially the TPR and the number of reconstructed states) show that the apparent contradiction between simulations and empirical application depend in all probability on the specific choice that was made in the simulation study, concerning structures and parameter values. A choice that could have facilitated, somehow Schrepp’s procedure.

Finally, it has to be stressed that all the obtained models present some identification problems. Such problems, as demonstrated in recent literature (Spoto et al., 2012, 2013),

Table 9 Binary representation of the skill map obtained for the empirical application with respect to the “Cleaning” sub-scale

Attributes						
Items	a	b	c	d	e	f
i_1	1	1	1	1	1	1
i_2	1	1	1	1	1	1
i_3	1	0	0	0	0	0
i_4	0	0	0	0	1	0
i_5	0	0	0	0	0	1
i_6	0	0	0	1	0	0
i_7	0	1	1	1	1	1
i_8	0	0	1	0	0	0

can be detected a priori from the obtained skill maps (when the proposed procedure is applied), or from the reconstructed structure.

With respect to the interpretability of the obtained skill maps, it emerged that two items investigate the same set of attributes, thus they can be considered as equivalent in conveying diagnostic information about cleaning rituals. Furthermore, a number of items investigate only one clinical attribute indicating that their information represents a specific element useful for a precise case formulation. Finally, item 7 provides information about a number of attributes that are investigated also by other items. This last issue could be fruitfully implemented into an automatic adaptive procedure that could allow for a reduction of the time needed to administer the scale.

General discussion

In this article, a procedure to construct a skill map starting from a data set has been described. This procedure is mainly based on the possibility of deriving a conjunctive skill map from a closure space or a disjunctive skill map from a knowledge space. The procedure begins, in the former case, with the minimal closure space (i.e., the one including only the empty set and the total set) and iteratively tries to improve such structure by minimizing the distance between the reconstructed closure space and the observed data. One of the most critical aspects of the procedure is the stopping criterion. In the present paper, the use of AICc has been tested and compared to the use of the *app* criterion developed by Schrepp for his procedure (Schrepp, 1999a) through a number of simulation studies. The results highlighted a better performance of this last index. In general, the procedure accurately reproduces the structure even when the error rate of the data (measured by parameters η and β) are relatively high (i.e., up to 0.30).

In order to test the applicability of the procedure in a real setting, two empirical studies were conducted on clinical data. These studies suggested the procedure is concretely applicable and that the obtained results are interpretable. It must be stressed that in this article, the procedure has been tested on relatively small structures. This was due to the high time consumption of the procedure to carry out the reconstruction on rather small domains. In order to apply the procedure to larger structures, the algorithm must be improved, since, in the present version, it appears to be too computationally expensive. Nevertheless, both the basic ideas and the method proposed in this study look very promising for a new way to build structures and skill maps starting from filled questionnaires.

The proposed procedure applies to conjunctive skill maps but can be trivially adapted to the disjunctive model. In that case, each structure in the chain will be closed under union, and the disjunctive skill map will simply correspond to the basis of the obtained space.

Lastly, the empirical application shows that the skills in the conjunctive skill map extracted from the data may then be interpreted by some experts in the field. This suggests the possibility of constructing hybrid procedures that rely on both the expert's opinion about the relationships between items and skills and the information contained in a data set. For instance, the expert may provide a rough initial skill map that is then further improved by an extraction method like the one presented in this article.

To summarize, the D-SMEP procedure presents a number of pros and some cons to be taken into account. The main cons are referred to the rather poor efficiency of the procedure in terms of computational time and to the stopping criterion. With respect to the former, it has to be stressed how a major improvement could be obtained by removing the closure constrain. In fact, if no closure is required, then the procedure may improve much in efficiency. Concerning the latter, further studies are needed. On the other hand, D-SMEP successfully copes with some of the main limitations of the already-existing procedures. First, the proposed procedure allows building an either conjunctive or disjunctive skill map starting from a set of observed data. D-SMEP could be modulated in order to obtain either a knowledge space or a closure space, or a quasi-ordinal knowledge space. Second, differently from Schrepp's procedure, it allows reconstructing states that are not observed in the sample at hand, provided that they are the intersection (or union) of observed states. Moreover, again in contrast with Schrepp's procedure, D-SMEP does not assume any equality constraints neither on the state probabilities nor on the η and β parameters. Finally, it proved to be robust even with high error rates.

References

- Bottesi, G., Spoto, A., Sanavio, E., & Vidotto, G. (2014). Beyond the score: clinical evaluation through formal psychological assessment. *Journal of Personality Assessment*. doi:10.1080/00223891.2014.958846.
- Brockwell, P. J., & Davis, R. A. (1991). *Time series: theory and methods*, 2nd edn. New York: Springer.
- Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference. A practical information-theoretic approach*, 2nd edn. New York: Springer.
- Cosyn, E., & Thiéry, N. (2000). A practical procedure to build a knowledge structure. *Journal of Mathematical Psychology*, 44, 383–407.

- DeLaTorre, J., & Douglas, J. (2008). Model evaluation and multiple strategies in cognitive diagnosis: an analysis of fraction subtraction data. *Psychometrika*, *73*, 595–624.
- Desmarais, M. C., & Maluf, A. (1996). User-expertise modeling with empirically derived probabilistic implication networks. *User Modeling and User-Adapted Interaction*, *5*, 283–315.
- Doignon, J. P. (1994). Knowledge spaces and skill assignments. In Fischer, G. H., & Laming, D. (Eds.) *Contributions to mathematical psychology, psychometrics, and methodology* (pp. 111–121). New York: Springer-Verlag.
- Doignon, J. P., & Falmagne, J. C. (1985). Spaces for the assessment of knowledge. *International Journal of Man-Machine Studies*, *23*, 175–196.
- Doignon, J. P., & Falmagne, J. C. (1999). *Knowledge spaces*. Berlin and Heidelberg: Springer-Verlag.
- Dütsch, I., & Gediga, G. (1995). Skills and knowledge structures. *British Journal of Mathematical Psychology*, *48*, 9–27.
- Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. Boca Raton, FL Chapman & Hall CRC.
- Falmagne, J. C., Doble, C., Albert, D., Eppstein, D., & Hu, X. (2013). *Knowledge spaces: Applications in education*. Berlin and Heidelberg: Springer-Verlag.
- Falmagne, J. C., & Doignon, J. P. (1988a). A class of stochastic procedures for the assessment of knowledge. *British Journal of Mathematical and Statistical Psychology*, *41*, 1–23.
- Falmagne, J. C., & Doignon, J. P. (1988b). A Markovian procedure for assessing the state of a system. *Journal of Mathematical Psychology*, *32*, 232–258.
- Falmagne, J. C., & Doignon, J. P. (2011). *Learning spaces*. Berlin and Heidelberg: Springer-Verlag.
- Falmagne, J. C., Koppen, M., Villano, M., Doignon, J. P., & Johannesen, L. (1990). Introduction to knowledge spaces: how to build, test and search them. *Psychological Review*, *97*, 201–224.
- Heller, J., Ünlü, A., & Albert, D. (2013). Skills, competencies and knowledge structures. In Falmagne, J. C., Doble, C., Albert, D., Eppstein, D., & Hu, X. (Eds.) *Knowledge spaces: Applications in education* (pp. 229–242). Berlin and Heidelberg: Springer-Verlag.
- Kambouri, M., Koppen, M., Villano, M., & Falmagne, J. C. (1993). Knowledge assessment: tapping human expertise by the QUERY routine. *International Journal of Human-Computer Studies*, *40*, 119–151.
- Koppen, M. (1993). Extracting human expertise for constructing knowledge spaces: an algorithm. *Journal of Mathematical Psychology*, *37*, 1–20.
- Koppen, M., & Doignon, J. P. (1990). How to build a knowledge space by querying an expert. *Journal of Mathematical Psychology*, *34*, 311–331.
- Korossy, K. (1997). Extending the theory of knowledge spaces: a competence-performance approach. *Zeitschrift für Psychologie*, *205*, 53–82.
- Sanavio, E., & Vidotto, G. (1985). The components of the Maudsley Obsessional-Compulsive Questionnaire. *Behaviour Research and Therapy*, *23*, 659–662.
- Sargin, A., & Ünlü, A. (2009). Inductive item tree analysis: corrections improvements and comparisons. *Mathematical Social Sciences*, *58*, 376–392.
- Schrepp, M. (1999a). Extracting knowledge structures from observed data. *British Journal of Mathematical and Statistical Psychology*, *52*, 213–224.
- Schrepp, M. (1999b). On the empirical construction of implications between bi-valued test items. *Mathematical Social Sciences*, *38*, 361–375.
- Schrepp, M. (2002). Explorative analysis of empirical data by Boolean analysis of questionnaires. *Zeitschrift für Psychologie*, *210*(2), 99–109.
- Schrepp, M. (2003). A method for the analysis of the hierarchical dependencies between items of a questionnaire. *Methods of Psychological Research - Online*, *19*, 43–79.
- Smith, N. A., & Tromble, R. W. (2004). Sampling uniformly from the unit simplex, Technical Report, John Hopkins University.
- Spoto, A., Bottesi, G., Sanavio, E., & Vidotto, G. (2013). Theoretical foundations and clinical implications of formal psychological assessment. *Psychotherapy and Psychosomatics*, *82*(3), 197–199.
- Spoto, A., Stefanutti, L., & Vidotto, G. (2010). Knowledge space theory, formal concept analysis, and computerized psychological assessment. *Behavior Research Methods*, *42*, 342–350.
- Spoto, A., Stefanutti, L., & Vidotto, G. (2012). On the unidentifiability of a certain class of skill map-based probabilistic knowledge structures. *Journal of Mathematical Psychology*, *56*(4), 248–255.
- Spoto, A., Stefanutti, L., & Vidotto, G. (2013). Considerations about the identification of forward- and backward-graded knowledge structures. *Journal of Mathematical Psychology*, *57*(4), 249–254.
- Stefanutti, L. (2006). A logistic approach to knowledge structures. *Journal of Mathematical Psychology*, *50*, 545–561.
- Stefanutti, L., & Koppen, M. (2003). A procedure for the incremental construction of a knowledge space. *Journal of Mathematical Psychology*, *47*, 265–277.
- Templin, J. L., & Henson, R. A. (2006). Measurement of psychological disorders using cognitive diagnosis models. *Psychological Methods*, *11*, 287–305.
- Ünlü, A., & Albert, D. (2004). The correlational agreement coefficient $CA(\leq, D)$ —a mathematical analysis of a descriptive goodness-of-fit measure. *Mathematical Social Sciences*, *48*, 281–314.
- Ünlü, A., Schrepp, M., Heller, J., Hockemeyer, C., Wesiak, G., & Albert, D. (2013). Recent developments in performance-based knowledge space theory. In Falmagne, J. C., Doble, C., Albert, D., Eppstein, D., & Hu, X. (Eds.) *Knowledge spaces: Applications in education* (pp. 147–192). Berlin and Heidelberg: Springer-Verlag.
- vanLeeuwe, J. F. J. (1974). Item tree analysis. *Nederlands Tijdschrift voor de Psychologie*, *29*, 475–484.