Available online at www.sciencedirect.com
SCIENCE DIRECT®
Mathematical Social Sciences 48 (2004) 281–314
mathematical social sciences
www.elsevier.com/locate/econbase

The Correlational Agreement Coefficient CA(≤,*D*)—a mathematical analysis of a descriptive goodness-of-fit measure

Ali Ünlü*, Dietrich Albert

Cognitive Science Section, Department of Psychology, University of Graz, Universitätsplatz 2/III, A-8010 Graz, Austria

Received 30 May 2003; received in revised form 25 January 2004; accepted 20 March 2004
Available online 20 June 2004

Abstract

The Correlational Agreement Coefficient, CA(≤,*D*), was introduced by J.F.J. van Leeuwe in 1974 within Item Tree Analysis (ITA), a data-analytic method to derive quasi orders (surmise relations) on sets of bi-valued test items. Recently, it has become of interest in connection with Knowledge Space Theory (KST). The coefficient CA(≤,*D*) is used as a descriptive goodness-of-fit measure to select out of competing surmise relations one with maximal CA(≤,*D*) value. Formal aspects like boundedness, decomposition, and the interplay between consistency of a surmise relation (with a binary data matrix) and the attainment of the maximum value of CA(≤,*D*) are investigated. Dependence of CA(≤,*D*) on trivial response patterns is quantified by a functional relationship that allows one to bunch the impact of trivial response patterns in a single "bias term". These considerations should warn against inconsiderate use of the coefficient. Mathematical reasons for failed, however, heuristically plausible, properties are presented.

Keywords: Correlational Agreement Coefficient; Item Tree Analysis; Knowledge Space Theory; Surmise relation; Descriptive goodness-of-fit measure

JEL classification: C00

*Corresponding author. Tel.: +43-316-380-5107; fax: +43-316-380-9806.
E-mail address: ali.uenlue@uni-graz.at (A. Ünlü).

0165-4896/$ - see front matter © 2004 All rights reserved.
doi:10.1016/j.mathsocsci.2004.03.002

1. Introduction

1.1. Preliminaries

In the field of knowledge assessment and acquisition based on prerequisite relationships, a central problem is to derive reflexive, transitive binary relations on sets of bi-valued test items. This is done for modeling hierarchies between items based on solvability dependencies of the type: "Given a positive response to an item J (e.g., J solved), it can be surmised that another item I will also be responded to positively (e.g., I solved)". Such binary relations (quasi orders) are central within Knowledge Space Theory (KST) introduced by Doignon and Falmagne (1985, 1999). In KST, they are called *surmise relations*. However, given a field of knowledge and a set of bi-valued test items appropriate enough to allow for fine-grained and representative coverage of the field, the problem is how to establish a reasonable surmise relation on the item set. Item Tree Analysis (ITA) is a data-analytic method for the derivation of surmise relations on sets of bi-valued test items. ITA was introduced by Airasian, Bart, and Krus in 1973 (Airasian and Bart, 1973; Bart and Krus, 1973) and was developed into the present form by Leeuwe (1974). In particular, Leeuwe (1974) introduced the Correlational Agreement Coefficient, CA($\leq$,D), as part of ITA.[1] In ITA, CA($\leq$,D) is used as a descriptive goodness-of-fit measure to select out of competing surmise relations one with maximal CA($\leq$,D) value.

Recently, ITA, and in particular, CA($\leq$,D), has become of interest in connection with KST; see Held et al. (1995), Held and Korossy (1998), Schrepp (1999, 2003), and Schrepp et al. (1999). For the application of efficient adaptive computer-based knowledge assessment procedures, one requires surmise relations of "a trade-off type". On the one hand, it should reflect the data as well as possible (*descriptive adequacy*) and on the other hand, it should be of as large as possible cardinality as a set. The authors tried to achieve this by applying ITA and the coefficient CA($\leq$,D) (cp. Section 12).

Leeuwe (1974) reports:[2] "*...This coefficient* [partial order reproducibility coefficient][3] *cannot serve therefore* [stationarity in tolerance level L=0] *as a criterion for choosing the best solution...This procedure* [CA($\leq$,D)] *has the advantage that it gives a lower value not only in the case that too many relations are constructed* [larger tolerance levels], *but also in the case that the number of relations is very low* [smaller tolerance levels]".

ITA's renaissance in connection with KST has led to criticisms of CA($\leq$,D). Held and Korossy (1998) stress the "ad hoc" (descriptive) nature of CA($\leq$,D): "*...we will apply two ad hoc criteria* [one, the CA($\leq$,D)]". Schrepp (1999) illustrates that CA($\leq$,D) can be reduced by non-comparable item pairs: "*...for relations which contain many non-connected item pairs it seems possible that the correct relation*

[1]The definition of CA($\leq$,D) will not be given until Section 6. For a coherent definition, the notions of empirical and theoretical correlation have to be introduced and compared. Till then, and this should not affect the elaborations to follow, the reader is asked to regard the coefficient as a real number resulting from an "input" surmise relation $\leq$ and binary data matrix D. It measures the goodness-of-fit of $\leq$ to D, and is practically interpreted as "the greater the value of CA($\leq$,D), the better the fit between $\leq$ and D".

[2]In the sequel, additional explanatory comments are put in square brackets.

[3]It is the proportion of subjects not contradicting the binary relation.

$\leq_L$ *will not have the best* CA($\leq_L$) *value*''. Another criticism of ITA and CA($\leq$,D) is voiced by Wesiak et al. (2004). They observe that trivial response patterns (i.e., all or none of the items answered positively), though empirically irrelevant with respect to solvability dependencies between items, do drastically manipulate ITA solutions. This is due to CA($\leq$,D)'s dependence on such patterns (cp. Section 11).

In the light of these observations, a comprehensive mathematical analysis of CA($\leq$,D) is missing. Rather, the elaborations so far are heuristic, based on experimentation with certain data sets. Other deeper properties of CA($\leq$,D) are actually not known so far. Thus, this work represents a coherent and extensive mathematical treatise on CA($\leq$,D). In particular, it warns against inconsiderate use of the coefficient, and if used, it tells to what one needs to pay attention. Perhaps, this work may also be viewed as a general guide to carry out a first mathematical analysis of ad hoc formulated coefficients. Additionally, Section 12 contains valuable methodological issues in regard to goodness-of-fit measures in general. Beside criteria proposed by Goodman and Kruskal (1954, 1959, 1963, 1972) (reviewed by Bishop et al., 1975; Liebetrau, 1983), Section 12 mentions the importance of purpose-specific goodness-of-fit measures and the problem of trade-off between different fit criteria.

1.2. A short review on ITA

This section reviews Leeuwe's (1974) Item Tree Analysis.

We use the following notation ($m,n \in \mathbb{N}$)[4]:

$Q := \{I_l \colon 1 \leq l \leq m\}$ set of dichotomous items,
$P := \{P_k \colon 1 \leq k \leq n\}$ sample of subjects,
$D := (d'_{kl})$ corresponding binary (=0/1) $n \times m$ data matrix,

and for every $(I_i, I_j) \in Q \times Q$ ($1 \leq i, j \leq m$), the 2×2 table notation

$\mathbf{I}_i \diagdown \mathbf{I}_j$	**1**	**0**
1	a_{ij}	b_{ij}
0	c_{ij}	d_{ij}

with $a_{ij}, b_{ij}, c_{ij}, d_{ij} \in \mathbb{N} \cup \{0\}$; in respective order, the absolute frequencies of subjects solving items I_i and I_j [a_{ij}], solving I_i, not I_j [b_{ij}], solving I_j, not I_i [c_{ij}], and solving neither I_i nor I_j [d_{ij}]. Then, the *ITA rule* for generating binary relations $\leq_L$ ($0 \leq L \leq n$) is given by

$$I_i \leq_L I_j \; :\Leftrightarrow c_{ij} \leq L.$$

[4] $\mathbb{N}$ stands for the set of natural numbers (excluding 0).

This L ($0 \le L \le n$) is called *tolerance level*. The ITA rule represents STEP1 of ITA. The latter consists of five steps, STEP1–STEP5:

STEP1. Determine the binary relations $\le_L$ for L=0, 1,. . ., n.
STEP2. From the $\le_L$ ($0 \le L \le n$), remove those that are not transitive.
STEP3. Set a critical value $0 < c \le 1$ for the proportions, p_L, of subjects not contradicting the respective surmise relations $\le_L$ in STEP2.
STEP4. From the surmise relations in STEP2, remove those with $p_L < c$.
STEP5. From the remaining surmise relations (after STEP4)—$\le_0$ is always contained—select one with maximal CA($\le$,D) value.

The Correlational Agreement Coefficient is used as a goodness-of-fit measure to handle the selection problem in STEP5. From the remaining surmise relations, select an "optimal" one, i.e., one with maximal CA($\le$,D) value.

1.3. On the content of this work

Basic concepts and the definition of empirical Pearson correlation are reviewed (Section 2). The definition of theoretical correlation is presented (Section 3). Empirical and theoretical correlation are compared in regard to coincidence (Section 4) and boundedness (Section 5). Based on this, CA($\le$,D) is defined coherently (Section 6). A natural decomposition of the coefficient CA($\le$,D) into four partial functions is given (Section 7). It is analyzed in regard to boundedness (Section 8). An analysis of the interplay between the consistency of a surmise relation $\le$ with a data matrix D and the attainment of the maximum value of CA($\le$,D) is presented (Sections 9 and 10). We conclude with the analysis of the dependence of CA($\le$,D) on trivial response patterns (Section 11). The work ends with a discussion (Section 12).

Note that all proofs are deferred to an appendix, section-wise (Appendix A).

2. Basic concepts

We review basic conventions regarding terminology and notation.

Let Q, P, and D be defined as in Section 1.2. The row z_k ($1 \le k \le n$) of D encodes the responses of subject P_k to all items in Q, whereas column s_l ($1 \le l \le m$) of D encodes the responses of all subjects in P to item I_l.

Definition 1. Let $Q=\{I_l\colon 1 \le l \le m\}$ ($m \in \mathbb{N}$). We define:[5]

$$\mathcal{S} := \{\le \subseteq Q \times Q : \le \text{ quasi order on } Q\},$$

$$\mathcal{D} := \bigcup_{n \in \mathbb{N}} \mathcal{M}(n \times m; \{0,1\}).$$

[5]Any reflexive, transitive binary relation is called quasi order (surmise relation). Let $\mathcal{M}(n \times m; \{0, 1\})$ denote the set of all $n \times m$ matrices with entries in $\{0,1\}$.

Let $(I_i,I_j)\in Q\times Q$. For the entries $a_{ij},b_{ij},c_{ij},d_{ij}\in\mathbb{N}\cup\{0\}$ of its 2×2 table (see Section 1.2), we have:

$$
\begin{aligned}
a_{ij} &= |\{k\in\{1,2,\ldots,n\} : d'_{ki}=1\wedge d'_{kj}=1\}|,\\
b_{ij} &= |\{k\in\{1,2,\ldots,n\} : d'_{ki}=1\wedge d'_{kj}=0\}|,\\
c_{ij} &= |\{k\in\{1,2,\ldots,n\} : d'_{ki}=0\wedge d'_{kj}=1\}|,\\
d_{ij} &= |\{k\in\{1,2,\ldots,n\} : d'_{ki}=0\wedge d'_{kj}=0\}|.
\end{aligned}
$$

Definition 2. For every $I_i\in Q$ $(1\le i\le m)$, let p_{I_i} be the relative frequency of subjects solving item I_i, i.e., $p_{I_i}:=n^{-1}\sum_{k=1}^{n} d'_{ki}$.

Definition 3. Let $\le$ be a surmise relation on Q. We shall say that $\le$ is consistent with (or, a total fit to) a data matrix D if, for every pair $(I_i,I_j)\in\le$, every subject in P solving item I_j also solves item I_i. In other words,

$$\forall (I_i,I_j)\in\le :\ |\{k\in\{1,2,\ldots,n\} : d'_{ki}=0\wedge d'_{kj}=1\}| = 0.$$

Empirical correlation is defined as sample Pearson correlation.

Definition 4. Let $I_i,I_j\in Q$ with $s_i\neq \underline{0}_n,\ \underline{1}_n$ and $s_j\neq \underline{0}_n,\ \underline{1}_n$.[6] Empirical correlation, r_{ij}, between I_i and I_j, given the observed data matrix D, is defined as (empirical) Pearson correlation between columns s_i and s_j of D:[7]

$$r_{ij}:=\frac{\mathrm{Cov}(s_i,s_j)}{\sqrt{\mathrm{Var}(s_i)}\sqrt{\mathrm{Var}(s_j)}}.$$

Definition 5. The set of all item pairs $(I_i,I_j)\in Q\times Q$, which fulfill the assumption $\mathbf{A}:=[s_i\neq\underline{0}_n,\ \underline{1}_n\wedge s_j\neq\underline{0}_n,\ \underline{1}_n]$, is denoted by $[Q\times Q]_{\mathbf{A}}$.

For the next two well-known results, see Bortz (1989).

Lemma 6. *For all $(I_i,I_j)\in[Q\times Q]_{\mathbf{A}}$, $r_{ij}\in[-1,1]$.*

[6] We have $\underline{0}_n:=(0,0,\ldots,0)^{\mathrm{T}}$ and $\underline{1}_n:=(1,1,\ldots,1)^{\mathrm{T}}$ (of length n).

[7] For $\mathbf{x}=(x_1,x_2,\ldots,x_n)^{\mathrm{T}},\mathbf{y}=(y_1,y_2,\ldots,y_n)^{\mathrm{T}}\in\mathbb{R}^n$ ($\bar{\mathbf{x}}$, $\bar{\mathbf{y}}$ empirical means):

$$\mathrm{Cov}(\mathbf{x},\mathbf{y}):=\frac{1}{n}\sum_{k=1}^{n}[(x_k-\bar{\mathbf{x}})(y_k-\bar{\mathbf{y}})],$$

$$\mathrm{Var}(\mathbf{x}):=\mathrm{Cov}(\mathbf{x},\mathbf{x}).$$

Lemma 7. *Let $(I_i,I_j)\in[Q\times Q]_A$ with corresponding 2×2 table*

$\mathbf{I}_i \diagdown \mathbf{I}_j$	**1**	**0**
1	a	b
0	c	d

Then, $r_{ij}=\dfrac{ad-bc}{\sqrt{(a+c)(b+d)(a+b)(c+d)}}$.

3. Theoretical correlation derived through idealization

Section 4 gives motivation for the form and name of theoretical correlation.

Definition 8. Let $(I_i,I_j)\in[Q\times Q]_{\mathbf{A}}$ and $\leq\in\mathcal{S}$. Theoretical correlation, r^*_{ij}, between I_i and I_j, derived through idealization, is defined as

$$r^*_{ij}:=\begin{cases} 1 & : (I_i,I_j)\in\leq \wedge (I_j,I_i)\in\leq \\ \sqrt{\dfrac{(1-p_{I_i})\cdot p_{I_j}}{(1-p_{I_j})\cdot p_{I_i}}} & : (I_i,I_j)\in\leq \wedge (I_j,I_i)\notin\leq \\ \sqrt{\dfrac{(1-p_{I_j})\cdot p_{I_i}}{(1-p_{I_i})\cdot p_{I_j}}} & : (I_i,I_j)\notin\leq \wedge (I_j,I_i)\in\leq \\ 0 & : (I_i,I_j)\notin\leq \wedge (I_j,I_i)\notin\leq \end{cases}$$

Theoretical correlation r^*_{ij} is well-defined for every $(I_i,I_j)\in[Q\times Q]_{\mathbf{A}}$. It is the case that $s_i,s_j\neq\underline{0}_n$, $\underline{1}_n$, i.e., p_{I_i}, $p_{I_j}\neq 0$, 1.

4. Comparing empirical and theoretical correlation: coincidence

Lemma 9. *Let $\leq\in\mathcal{S}$, which is consistent with binary response data D. Then, for all $(I_i,I_j)\in\leq\cap\,[Q\times Q]_A$,*

$$r_{ij}=\begin{cases} 1 & : (I_j,I_i)\in\leq \\ \sqrt{\dfrac{(1-p_{I_i})p_{I_j}}{(1-p_{I_j})p_{I_i}}} & : (I_j,I_i)\notin\leq \end{cases}$$

Proof. See Appendix A.1. □

The next corollary gives a first answer to the question of coincidence.

Corollary 10. *Let $\leq\in\mathcal{S}$, consistent with D. Let $(I_i,I_j)\in\leq$ with $\boldsymbol{A}$. Then, theoretical correlation r^*_{ij} equals empirical correlation r_{ij} (i.e., $r^*_{ij}=r_{ij}$).*

What can be said about coincidence in case of not-$\leq$-comparable item pairs?[8] The counterexample in Lemma 11 goes back to Schrepp (1999, pp. 364–365).

Lemma 11 (Counterexample). *Let $\leq$ be a surmise relation on Q, which is consistent with D. Let $(I_i,I_j)\in[Q\times Q]_A$ be a not-$\leq$-comparable item pair. Then, i.g.,*[9] $r_{ij}\neq 0$ $(=:r^*_{ij})$.

Proof. See Appendix A.1. □

In order to sum up, we give Proposition 12.

Proposition 12. *Let $\leq\in\mathcal{S}$, consistent with D. Let $(I_i,I_j)\in[Q\times Q]_A$. For $\delta_{ij}:=r_{ij}-r^*_{ij}$,*

$$\delta_{ij}\begin{cases} =0 & : (I_i,I_j)\in\leq \vee (I_j,I_i)\in\leq \\ \neq 0 \ i.g. & : (I_i,I_j)\notin\leq \wedge (I_j,I_i)\notin\leq \end{cases}$$

Proof. See Appendix A.1. □

4.1. Contrary properties

The exception with respect to coincidence described in Proposition 12 diminishes CA($\leq$,D)'s plausibility as a reasonable goodness-of-fit measure:

1. Whereas the definition of r^*_{ij} could be motivated in the first three cases, $r^*_{ij}:=0$ for not-$\leq$-comparable pairs lacks a clear statistical motivation. In general, $r^*_{ij}:=0$ is not the Pearson correlation for not-$\leq$-comparable pairs resulting from $\leq$-consistent data matrices. However, it can be shown that $r^*_{ij}:=0$ is the arithmetic mean of all such Pearson correlation values. An open question is whether 0 is the most frequently appearing value in the frequency distribution of these Pearson correlation values.
2. The consistency of a surmise relation $\leq$ with a data matrix D may not imply the attainment of the maximum value (cp. Section 8) CA($\leq$,D)=1.[10] This is only due to not-$\leq$-comparable item pairs, which is discussed in Section 9. For such pairs, theoretical correlation, i.g., differs from empirical correlation, and therefore CA($\leq$,D) is reduced.
3. In ITA, selecting an "optimal" surmise relation is based on the maximal value of CA($\leq$,D). As to point 2 above, not-$\leq$-comparable pairs may reduce CA($\leq$,D), even in case of a total fit. Thus, it is conceivable that a suitable (maybe correct) ITA structure, requiring a lot of not-$\leq$-comparable pairs, is ruled out from selection because of a distortedly low CA($\leq$,D) value (cp. Schrepp, 1999).

5. Comparing empirical and theoretical correlation: boundedness

Empirical correlation uniformly lies in the interval $[-1,1]$ (Lemma 6). What about theoretical correlation?

[8] A pair $(I_i,I_j)\in Q\times Q$ is called not-$\leq$-comparable iff $(I_i,I_j)\notin\leq$ and $(I_j,I_i)\notin\leq$.
[9] In this work, i.g. stands for "in general".
[10] In this work, we use the word "maximum" to refer to the attainment of the maximum value CA($\leq$,D)=1.

Proposition 13. *Let $Q=\{I_l: 1\leq l\leq m\}$ $(m\in\mathbb{N})$. It holds:*

***(Relative Interval Nesting)**. Let $D\in\mathcal{M}(n\times m; \{0,1\})$, $\leq\in\mathcal{S}$, and let $(I_i,I_j)\in[Q\times Q]_{A}$. Then (since $(I_i,I_j)\in[Q\times Q]_{A}$, $n\geq 2$),*

$$0\leq r^*_{ij}\leq n-1.$$

***(Proper Divergence to $+\infty$)**. For $m\geq 2$, there exists an $\leq_*\in\mathcal{S}$ and a pair $(I_i,I_j)\in Q\times Q$ with $i<j$ and $[(I_i,I_j)\in\leq_*\wedge(I_j,I_i)\notin\leq_*]$, such that*

$$\forall n\geq 2\,\exists D_{n-1}\in\mathcal{M}(n\times m;\{0,1\}):[[(r_{ij})_{n-1}=-1]\wedge[(r^*_{ij})_{n-1}=n-1]].$$ [11]

The corresponding sequence $((r^*_{ij})_n)_{n\in\mathbb{N}}$ diverges properly to $+\infty$,[12]

$$\lim_{n\to\infty}((r^*_{ij})_n)_{n\in\mathbb{N}}=+\infty.$$

Proof. See Appendix A.2. □

5.1. Informal illustration of (PDt+ ∞)

Let $Q=\{I_1,I_2,I_3,I_4\}$, and $\leq:=\Delta\cup\{(I_1,I_2)\}$.[13] Consider the scheme:[14]

$$D_1:=\begin{bmatrix}1&0&0&0\\0&1&0&0\end{bmatrix}_{2\times 4}\quad\rightarrow\quad(r^*_{12})_1=1,$$

$$D_2:=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&1&0&0\end{bmatrix}_{3\times 4}\quad\rightarrow\quad(r^*_{12})_2=2,$$

$$D_3:=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&1&0&0\\0&1&1&1\end{bmatrix}_{4\times 4}\quad\rightarrow\quad(r^*_{12})_3=3,$$

$$\vdots\qquad\downarrow\qquad+\infty.$$

[11]The symbols $(r_{ij})_{n-1}$ and $(r^*_{ij})_{n-1}$ $(n\geq 2)$ respectively denote empirical and theoretical correlation between I_i and I_j, given the data matrix D_{n-1}.

[12]A sequence $(a_n)_{n\in\mathbb{N}}$ of real numbers *diverges properly* to $+\infty$ (respectively $-\infty$) if, for every $K\in\mathbb{R}$, there is an $N\in\mathbb{N}$, such that $a_n>K$ (respectively $a_n<K$) for all $n\geq N$. In this case, we write $\lim_{n\to\infty}(a_n)_{n\in\mathbb{N}}=+\infty$ (respectively $\lim_{n\to\infty}(a_n)_{n\in\mathbb{N}}=-\infty$).

[13]We have $\Delta:=\{J=(J_1,J_2)\in Q\times Q: J_1=J_2\}$.

[14]Note that $(r_{12})_1=-1$, $(r_{12})_2=-1$, $(r_{12})_3=-1$, ….

5.2. Contrary properties

There is no uniform, but rather relative interval nesting for CA($\leq$,D) values, depending on the number n of rows of the data matrix D (see Section 8). Contrary to psychological heuristics, CA($\leq$,D) ranging in the interval [0,1], the latter may take negative values, even diverge properly to $-\infty$. The reason for this is the difference between empirical and theoretical correlation in regard to boundedness. Whereas empirical correlation is uniformly bounded by [−1,1], there is no such uniform interval nesting for theoretical correlation.

The next corollary is a direct consequence of Proposition 13.

Corollary 14. *Let $\leq$ be a surmise relation on Q, and let $(I_i,I_j)\in[Q\times Q]_{\mathbf{A}}$. For $\delta_{ij}:=r_{ij}-r^*_{ij}$, we have $-n\leq\delta_{ij}\leq 1$. In particular, $0\leq\delta^2_{ij}\leq n^2$.*

Proof. See Appendix A.2. □

6. Defining the coefficient CA($\leq$,D)

Definition 15. Let $Q:=\{I_l\colon 1\leq l\leq m\}$ ($m\in\mathbb{N}$, $m\geq 2$), $\leq$ be a surmise relation on Q, and $D=(d'_{kl})\in\mathcal{M}(n\times m;\ \{0,1\})$. Further, let

$$<'_Q:=\{(I_i,I_j)\in Q\times Q : i<j \text{ and } (I_i,I_j) \text{ fulfills } \mathbf{A}\}.$$

The **Correlational Agreement Coefficient**, CA($\leq$,D), is defined as

$$\mathrm{CA}(\leq,D):=1-\frac{2}{m(m-1)}\sum_{(I_i,I_j)\in<'_Q}(r_{ij}-r^*_{ij})^2.$$

We close this section with two (actually obvious) remarks.

6.1. (Artificial) convention

For $<'_Q=\emptyset$ (e.g., n=1 implies $<'_Q=\emptyset$), we agree upon

$$\sum_{(I_i,I_j)\in\emptyset}(r_{ij}-r^*_{ij})^2:=0.$$

Then, $\mathrm{CA}(\leq,D)=1-2[m(m-1)]^{-1}0=1$. Independent of the collection $\leq$ of solvability dependencies, we have (in case of $<'_Q=\emptyset$) CA($\leq$,D)=1.

6.2. Conception as a function[15]

The Correlational Agreement Coefficient can be interpreted as a function of domain $\mathcal{S}\times\mathcal{D}$ and co-domain $\mathbb{R}$,

$$\mathrm{CA}:\mathcal{S}\times\mathcal{D}\to\mathbb{R},\ (\leq,D)\mapsto\mathrm{CA}((\leq,D)):=\mathrm{CA}(\leq,D).$$

[15]Note that $\mathcal{S}:=\{\leq\subseteq Q\times Q\colon \leq \text{ quasi order on Q}\}$, $\mathcal{D}:=\bigcup_{n\in\mathbb{N}}\mathcal{M}(n\times m;\ \{0,1\})$.

7. Decomposing the coefficient CA($\leq$,D)

We begin with some notation.

Definition 16. Let Q:={I_l: 1≤l≤m} (m≥2), D∈$\mathcal{M}$(n×m; {0,1}), and ≤∈$\mathcal{S}$. We define:

$$\begin{aligned}
<'_{Q|\cong} &:= <'_Q \cap \{(I_i,I_j)\in Q\times Q : (I_i,I_j)\in\leq \wedge (I_j,I_i)\in\leq\},\\
<'_{Q|\ll} &:= <'_Q \cap \{(I_i,I_j)\in Q\times Q : (I_i,I_j)\in\leq \wedge (I_j,I_i)\notin\leq\},\\
<'_{Q|\gg} &:= <'_Q \cap \{(I_i,I_j)\in Q\times Q : (I_i,I_j)\notin\leq \wedge (I_j,I_i)\in\leq\},\\
<'_{Q|\not\cong} &:= <'_Q \cap \{(I_i,I_j)\in Q\times Q : (I_i,I_j)\notin\leq \wedge (I_j,I_i)\notin\leq\}.
\end{aligned}$$

The family $\mathcal{F}$:=($<'_{Q|\cong}$,$<'_{Q|\ll}$,$<'_{Q|\gg}$,$<'_{Q|\not\cong}$) of subsets of $<'_Q$ fulfills

$<'_Q = \bigcup_{k\in\{\cong,\ll,\gg,\not\cong\}} <'_{Q|k}$ (Coverging property),

$<'_{Q|k} \cap <'_{Q|l} = \emptyset$ for $k,l\in\{\cong,\ll,\gg,\not\cong\}, k\neq l$ (Pairwise disjoint).

In general, $\mathcal{F}$ may not be a partition of $<'_Q$, since one of the members $<'_{Q|i}$ could be empty. In other words, $\mathcal{F}$ constitutes a disjoint covering of $<'_Q$:

$$<'_Q = \sum_{k\in\{\cong,\ll,\gg,\not\cong\}} <'_{Q|k}.$$

Next, we give the natural decomposition of CA(≤,D).

Lemma 17 (Decomposition). *Let Q={I_l: 1≤l≤m} (m∈ℕ, m≥2), $\mathcal{D}$∈$\mathcal{M}$(n×m; {0,1}), and ≤∈S. Then, the value CA((≤,D)):=CA(≤,D) can be decomposed in the following way:*

$$CA(\leq,D) = 1 - \frac{2}{m(m-1)}\sum_{k=1}^{4} f_k(\leq,D),$$

whereupon the values f_k(≤,D)∈ℝ (1≤k≤4) are defined as[16] *(with $Z:=[a_{ij}+b_{ij}]\cdot[c_{ij}+d_{ij}]\cdot[a_{ij}+c_{ij}]\cdot[b_{ij}+d_{ij}]$)*

$$f_1(\leq,D) := \sum_{(I_i,I_j)\in<'_{Q|\cong}} \left[\frac{a_{ij}d_{ij}-b_{ij}c_{ij}}{\sqrt{Z}} - 1\right]^2,$$

$$f_2(\leq,D) := \sum_{(I_i,I_j)\in<'_{Q|\ll}} \frac{[nc_{ij}]^2}{Z},$$

$$f_3(\leq,D) := \sum_{(I_i,I_j)\in<'_{Q|\gg}} \frac{[nb_{ij}]^2}{Z},$$

$$f_4(\leq,D) := \sum_{(I_i,I_j)\in<'_{Q|\not\cong}} \frac{[a_{ij}d_{ij}-b_{ij}c_{ij}]^2}{Z}.$$

In particular, f_k: $\mathcal{S}\times\mathcal{D}\to\mathbb{R}$, $(\leq,D)\mapsto f_k((\leq,D)):=f_k(\leq,D)$, 1≤k≤4.

[16]Remember that $a_{ij},b_{ij},c_{ij},d_{ij}\in\mathbb{N}\cup\{0\}$ are denoting the entries in the 2×2 table for a pair $(I_i,I_j)\in Q\times Q$ (Section 1.2). Note that $\sum_{(I_i,I_j)\in\emptyset}\cdots:=0$.

Proof. See Appendix A.3. □

8. Boundedness of CA($\leq$,*D*)

Proposition 18. *Let $Q:=\{I_l: 1\leq l\leq m\}$ ($m\geq 2$). It holds:*

(Relative Interval Nesting)*. If $D\in\mathcal{M}(n\times m;\ \{0,1\})$ for $n\in\mathbb{N}$ fixed, then, for all $\leq\in\mathcal{S}$,*

$$1-n^2\leq CA(\leq,D)\leq 1.$$

That is, partial function $CA(.,D)$: $\mathcal{S}\to\mathbb{R}$, $\leq\mapsto CA(.,D)(\leq):=CA(\leq,D)$ has a bounded range $CA(.,D)(\mathcal{S})\subset[1-n^2,1]$.

(Proper Divergence to $-\infty$)*. There exists an $\leq_*\in\mathcal{S}$ and $(D_n)_{n\in\mathbb{N}}$ in $\mathcal{D}$:*

$$\lim_{n\to\infty}(CA(\leq_*,D_n))_{n\in\mathbb{N}}=-\infty,$$

in the sense of diverging properly to $-\infty$.

Proof. See Appendix A.4. □

8.1. Informal illustration of (PDt$-\infty$)

Reconsider the scheme on p. 8 of this article:

$$D_1:=\begin{bmatrix}1&0&0&0\\0&1&0&0\end{bmatrix}_{2\times4}\quad\to\quad \mathrm{CA}(\leq,D_1)=+0.33,$$

$$D_2:=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&1&0&0\end{bmatrix}_{3\times4}\quad\to\quad \mathrm{CA}(\leq,D_2)=-0.50,$$

$$D_3:=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&1&0&0\\0&1&1&1\end{bmatrix}_{4\times4}\quad\to\quad \mathrm{CA}(\leq,D_3)=-1.91,$$

$$\vdots$$

$$\downarrow$$

$$-\infty.$$

Remark. The measure CA($\leq$,*D*) depends on the squared difference $(r_{ij}-r^*_{ij})^2\to+\infty$:

$$\mathrm{CA}(\leq,D):=\underbrace{1-\underbrace{\frac{2}{m(m-1)}\sum_{(I_i,I_j)\in<'_Q}\underbrace{(r_{ij}-r^*_{ij})^2}_{\to+\infty}}_{\to+\infty}}_{\to-\infty}\ .$$

This should not be viewed as an inadequacy that diminishes CA($\leq$,D)'s plausibility as a reasonable goodness-of-fit measure. It is contrary to convention and preconception that a "convenient" measure should range in [0,1] or [−1,1].[17] However, for the purpose of selecting between competing surmise relations, such a "standardization" is not relevant. Rather, it is important that the coefficient-based selection criterion is interpretable (meaningful). Since CA($\leq$,D) lacks a clear interpretation (cp. Section 12), this convention might seem to be important. However, the importance diminishes as the interpretability (meaningfulness) of a measure increases (Goodman and Kruskal, 1954).[18]

Or, if we speak in terms of functions:

Corollary 19. *Let* $Q{:=}\{I_l{:}\ 1\leq l\leq m\}$ $(m\geq 2)$ *and* $\mathcal{D}\in\mathcal{M}(n\times m;\ \{0,1\})$. *The partial function* $CA(.,D){:}\ \mathcal{S}\rightarrow\mathbb{R},\ \leq\mapsto CA(.,D)(\leq){:=}CA(\leq,D)$ *has bounded range* $CA(.,D)(\mathcal{S})\subset[1-n^2,1]$. *For the range* $CA(\mathcal{S}\times\mathcal{D})$ *of* $CA{:}\ \mathcal{S}\times\mathcal{D}\rightarrow\mathbb{R}$, *we have* $CA(\mathcal{S}\times\mathcal{D})\subset]-\infty,1]$,[19] *whereupon there is no* $[a,b]\subset\mathbb{R}$ *with* $CA(\mathcal{S}\times\mathcal{D})\subset[a,b]$.

9. Consistency–maximum problem

Reconsider the example in Lemma 11:

Lemma 20 (Counterexample). *Let* $Q{:=}\{I_l{:}\ 1\leq l\leq m\}$ $(m\geq 2)$ *and* $\leq\in\mathcal{S}$ *be a total fit to D. Then, it is not necessarily the case that* $CA(\leq,D)=1$. *In other words, consistency does not imply maximum in general.*

Proof. See Appendix A.5. □

Remark. If we presuppose consistency, and that CA($\leq$,D) depends on $(r_{ij}-r^*_{ij})^2>0$ for a not-$\leq$-comparable item pair $(I_i,I_j)\in<'_Q$, then we have:

$$\begin{aligned}
\mathrm{CA}(\leq,D) &{:=} 1-\frac{2}{m(m-1)}\sum_{(I_i,I_j)\in<'_Q}(r_{ij}-r^*_{ij})^2\\
&= \underbrace{1-\frac{2}{m(m-1)}\underbrace{\sum_{(I_i,I_j)\in<'_Q|_{\neq},\delta_{ij}>0}\underbrace{(r_{ij}-r^*_{ij})^2}_{>0}}_{>0}}_{<1}.
\end{aligned}$$

Equivalence between consistency and maximum is not a property of essential relevance for the purpose of selecting between competing surmise relations. The measure only needs to be meaningful. Meaningfulness presupposed, such an additional property would allow for a measure assuming its extreme value(s) for a certain type of "complete" association. For example, in case of CA($\leq$,D), this would mean the attainment of the maximum value 1 for the type of complete association "$\leq$ a total fit to D".

[17] A reason might be the fact that most of the traditional association measures, based on Pearson's χ^2 statistic (Pearson, 1904), are "standardized"; e.g., Pearson's coefficient of contingency (Pearson, 1904), Tschuprow's contingency coefficient (Tschuprow, 1918/1919/1921), and Cramér's contingency coefficient (Cramér, 1946).

[18] For traditional measures of association, see also Yule and Kendall (1950), Kendall and Stuart (1973), Goodman and Kruskal (1954, 1959, 1963, 1972), Bishop et al. (1975), and Liebetrau (1983).

[19] We have $]-\infty,1]{:=}\{x\in\mathbb{R}{:}\ x\leq 1\}$.

Corollary 21. *Let* $Q:=\{I_l: 1\leq l\leq m\}$ $(m\geq 2)$ *and* $\leq\in\mathcal{S}$ *be consistent with D. Then, it holds:*

1. *If* $<'_{Q|\neq}=\emptyset$, *then* $CA(\leq,D)=1$.
2. *If for every* $(I_i,I_j)\in<'_{Q|\neq}$, $r_{ij}=0$, *then* $CA(\leq,D)=1$.
3. *Even more,* $CA(\leq,D)=1 \Leftrightarrow \forall(I_i,I_j)\in<'_{Q|\neq}: r_{ij}=0$.

10. Maximum−consistency problem

The converse implication is also not true in general.

Lemma 22 (Counterexample). *Let* $Q:=\{I_l: 1\leq l\leq m\}$ $(m\geq 2)$, $D\in\mathcal{D}$, *and* $\leq\in\mathcal{S}$ *with* $CA(\leq,D)=1$. *Then, it is not necessarily the case that* $\leq$ *is consistent with D. In other words, maximum does not imply consistency in general.*

Proof. See Appendix A.6. □

Proposition 23 states that maximum CA($\leq$,D)=1 implies consistency, provided no subject contradicts any of the non-reflexive[20] pairs $I_i\leq I_j$ with non-existent empirical correlation r_{ij}.

Proposition 23. *Let* $Q:=\{I_l: 1\leq l\leq m\}$ $(m\geq 2)$. *Further let* $\leq\in\mathcal{S}$ *and* $D=(d'_{kl})\in\mathcal{M}(n\times m;\{0,1\})$ *such that* $CA(\leq,D)=1$. *If for every non-reflexive pair* $(I_i,I_j)\in\leq$ *with* $\neg\mathbf{A}$,[21] the entry c_{ij} *of its* 2×2 *table,*

$$c_{ij}(=|\{k\in\{1,2,\ldots,n\}: d'_{kji}=0 \wedge d'_{kj}=1\}|)=0,$$

then $\leq$ *is a total fit to D.*

Proof. See Appendix A.6. □

10.1. Contrary properties

Contrary to psychological heuristics, CA($\leq$,D) may attain its maximum 1 without $\leq$ being a total fit to D. This is only due to non-reflexive pairs $(I_i,I_j)\in\leq$ with non-existent empirical correlation r_{ij}.[22] Based on this, one can construct an example which consists of a "highly" inconsistent surmise relation $\leq_1$ with CA($\leq_1$,D)=1, and a total fit $\leq_2$ with CA($\leq_2$,D)<1. This might be regarded an inadequacy of CA($\leq$,D)'s plausibility as a reasonable selection criterion in ITA.

11. Functional relationship for equivalent data matrices

Wesiak et al. (2004) observe a "data-related" problem arising when trivial response patterns are included/excluded in/from the input data matrix for ITA. Such response patterns,

[20]Any pair of items $(J_1,J_2)\in Q\times Q$ with $J_1\neq J_2$ is called non-reflexive.
[21]$\neg$**A** stands for the negation, "not **A**", of **A**.
[22]Compare the counterexample in Lemma 22: Contradiction is only due to the non-reflexive pair $(I_3,I_2)\in\leq$ with non-existent empirical correlation r_{32}.

though empirically irrelevant with respect to solvability dependencies between items, do drastically manipulate ITA solutions. Larger/smaller optimal L_{opt} (stronger/weaker structures $\leq_{opt}$) are obtained by adding/removing trivial patterns to/from the input data.

Lemma 24 bunches the impact of such patterns in a single "bias term".

Lemma 24 (Functional relationship). *Let* $Q{:=}\{I_l{:}\ 1\leq l\leq m\}$ $(m\geq 2)$, $\leq\in\mathcal{S}$, *and* $D\in\mathcal{M}(n\times m;\{0,1\})$. *For* $n_0,n_1\in\mathbb{N}\cup\{0\}$, *let* D^t *represent D enriched with* n_0 *empty response patterns* $(0,0,\ldots,0)$ *and* n_1 *full response patterns* $(1,1,\ldots,1)$ *(i.e., as matrix rows). Under these conditions, with a special function* $F^*{:}\ (\mathbb{N}\cup\{0\})\times(\mathbb{N}\cup\{0\})\times\mathcal{D}\times\mathcal{S}\to\mathbb{R}$,

$$CA(\leq,D^t) = CA(\leq,D) - F^*(n_0,n_1,D,\leq).$$

Proof. See Appendix A.7. □

12. Discussion

12.1. Major misconceptions in CA(≤,D) publications

Two major misconceptions are present in some of the CA($\leq$,D) publications mentioned in Section 1.1:

(A) The coefficient CA($\leq$,D) does not measure goodness-of-fit with respect to the fit criterion "number of response patterns in D matching all pairs in $\leq$". In the terminology of knowledge spaces (Doignon and Falmagne, 1999), this is refered to as "number of response patterns in D matching one of the knowledge states in the quasi ordinal knowledge space $\mathcal{K}_{\leq}$, corresponding to $\leq$".[23]

(B) The coefficient CA($\leq$,D) does not provide a trade-off[24] between the goodness-of-fit of $\leq$ to D and size $|\leq|$ of $\leq$. In the terminology of knowledge spaces, this is refered to as a trade-off between goodness-of-fit and size of $\mathcal{K}_{\leq}$.[25]

In the literature, we find comments such as:

Held and Korossy (1998): "... The CA($\leq$,D) value takes into account the trade-off between absolute goodness of fit... and the total number of inferred knowledge states". Leeuwe (1974): "...This procedure [CA($\leq$,D)] has the advantage that it gives a lower value not only in the case that too many relations are constructed, but also that the number of relations is very low".

[23]This is Birkhoff's (1937) theorem. It states that "there is a one-to-one correspondence between the collection of all surmise relations on an item set Q, and the collection of all quasi ordinal knowledge spaces on Q". See Doignon and Falmagne (1999).

[24]Trade-off is understood in an informal way as "appropriate averaging".

[25]Whether goodness-of-fit is based on the fit criterion "match of response patterns in the structure", or (actually underlying CA($\leq$,D)) "agreement between empirical and theoretical correlation", the measure CA($\leq$,D) does not provide a trade-off.

Held et al. (1995): "... Mit Hilfe von DA wird ebenso wie bei CA der Trade-Off zwischen Größe und Passung einer Wissensstruktur berücksichtigt".[26]

Why do (A) and (B) occur? Ad[27] (A). The actual fit criterion underlying CA($\leq$,D) is "agreement between empirical and theoretical correlation" (FC2). Let FC1 denote the fit criterion "match of response patterns in the structure". Then, FC2 association is not "directionally informative" for FC1 association.[28] Consistency (i.e., "complete" FC1 association) does not imply maximum (i.e., "complete" FC2 association), and vice versa (Sections 9 and 10). This also holds for intermediate stages of degree of association. For instance, CA($\leq_1$,D)<CA($\leq_2$,D) ("$\leq_1$ smaller in degree of FC2 association than $\leq_2$") does not imply a directional statement of type "$\leq_1$ smaller in degree of FC1 association than $\leq_2$". Rather, there are cases with the reversed order of degree of FC1 association.

Ad (B). A counterexample is given by Wesiak et al. (2004). For an empirical data set, they obtain an optimal tolerance level $L_{\text{opt}}=0$. The "optimal" structure $\leq_0$ is a total fit to D ("complete" FC1 association). It maximizes CA($\leq$,D), i.e., in degree of FC2 association, $\leq_0$ is larger than or equal to any of the competing surmise relations. However, $\leq_0$ is not appropriately sized. It consists of 80 pairs of altogether $27^2=729$ pairs ($|Q|=27$; circa 11%). This structure is certainly too weak. An informal "mathematical" argument for (B) would be: In the definition of CA($\leq$,D), summation is taken over the "universal set"

$$<_Q^{\prime}:=\{(I_i,I_j)\in Q\times Q : i<j \text{ and } (I_i,I_j) \text{ fulfills } \mathbf{A}\},$$

the same set for all surmise relations considered. Standardization is with respect to the "universal size" $m(m-1)/2$, independent of size $|\leq|$. Thus, there is no "structure-specific averaging", which might give a hint for trade-off.

12.2. Major results in this paper

This work includes five major results:

1. *Decomposition* (Section 7). The coefficient CA($\leq$,D) is naturally decomposed into four finer parts. The decomposition is used to handle the maximum–consistency problem (Section 10) and the functional relationship (Section 11) in an elegant way. It also provides a useful descriptive "screening" method (exemplified below).
2. *Boundedness* (Section 8). The coefficient CA($\leq$,D) may take on negative values, even diverge properly to $-\infty$ (Proposition 18). The reason is that there is a difference between empirical and theoretical correlation in regard to boundedness (Section 5).
3. *Consistency–maximum problem* (Section 9). Consistency does not imply maximum in general. The reason is that there is a difference between empirical and theoretical

[26] In English: "By means of DA [another descriptive measure], just like CA($\leq$,D), the trade-off between size and goodness-of-fit of a knowledge structure is taken into account".

[27] In this work, the Latin word "ad" is used to stand for "in regards to".

[28] That is, higher degree of FC2 association implying higher degree of FC1 association.

correlation in regard to coincidence (Section 4). There are conditions under which the implication holds (Corollary 21).

4. *Maximum–consistency problem* (Section 10). Maximum does not imply consistency in general. This is only due to non-reflexive pairs of a surmise relation for which the empirical correlation does not exist. There is a condition under which the implication holds (Proposition 23).
5. *Functional relationship for equivalent data matrices* (Section 11). The functional relationship bunches the impact of trivial response patterns on CA($\leq$,D) in a single "bias term" (Lemma 24).

What do these results imply for applications of CA($\leq$,D)?

The coefficient CA($\leq$,D) is an ad hoc measure lacking a clear statistical foundation. It is purely used descriptive, no sampling and inference are considered. Thus, currently, we suggest that CA($\leq$,D) should be regarded as a tool for exploratory data and structure analysis (see Section 12.3). Nevertheless, the user should consider the following hints when using CA($\leq$,D):

Ad 2: In most of the empirical studies, CA($\leq$,D)$\in[0,1]$. However, it is the case that $1-n^2\leq$CA($\leq$,D)≤ 1 with even CA($\leq$,D)<0.

Ad 3: Do not interpret CA($\leq$,D)<1 with non-consistency. To assure consistency implying maximum, check for the conditions in Corollary 21.

Ad 4: Be aware that CA($\leq$,D) can attain its maximum 1 though $\leq$ may contradict the data D. To assure maximum implying consistency, check for the condition in Proposition 23. If you are using CA($\leq$,D) in the context of ITA, remove all $I_i\in Q$ with $s_i = \underline{1}_n$ or $s_i = \underline{0}_n$ before applying ITA.

Ad 5: Remove trivial response patterns from the data at the beginning of data analysis, and before calculating and utilizing CA($\leq$,D) within ITA.

12.3. Consequences of the decomposition on screening of data and structure

12.3.1. Properties of the functions f_i

Let f_i ($1\leq i\leq 4$) be the partial functions in the decomposition of CA($\leq$,D) (Lemma 17). They fulfill the following properties:[29,30]

(1) Obviously, for all $\leq\in\mathcal{S}$ and $D\in\mathcal{D}$, $f_i(\leq,D)\geq 0$ ($1\leq i\leq 4$).

(2) For $Q{:=}\{I_l\text{: } 1\leq l\leq m\}$ ($m\geq 2$) fixed, there is a global interval nesting for f_1 and f_4; e.g., $f_1(\mathcal{S}\times\mathcal{D})\subset[0,4m^2]$ and $f_4(\mathcal{S}\times\mathcal{D})\subset[0,m^2]$. But there is no such global interval nesting for f_2 and f_3 (Q fixed). The latter may diverge properly to $+\infty$. However, additionally fixing a matrix $D\in\mathcal{M}(n\times m;\ \{0,1\})$, there is a relative interval nesting for f_2 and f_3, depending on the number n of rows of D (cp. Corollary 14); e.g., $f_j(\mathcal{S}\times\mathcal{D})\subset[0,m^2n^2]$ for j=2,3.

[29]No proofs are given. These can be infered in analogy to the preceding proofs.

[30]The question is: How do the properties boundedness, consistency–maximum problem, maximum–consistency problem, and functional relationship, already analyzed for the global measure CA($\leq$,D), look for the local measures f_i?

(3) Local consistency is introduced for the sets $<'_{Q|k}$ ($k \in \{\cong, \ll, \gg\}$) in the decomposition of CA($\leq$,D):[31]

Definition 25. Let Q:=$\{I_l$: $1 \leq l \leq m\}$ ($m \geq 2$), $D \in \mathcal{M}(n \times m;\ \{0,1\})$, and $\leq \in$S. $<'_{Q|\cong}$ is called locally consistent with (or, local total fit to) D if no response pattern in D contradicts any of the pairs in $\leq$ determined by any of the elements in $<'_{Q|\cong}$. In other words, if

$$\forall (I_i, I_j) \in <'_{Q|\cong}: [I_i \leq I_j \text{ and } I_j \leq I_i \text{ do not contradict } D].$$

(Global) consistency implies local consistency of $<'_{Q|k}$ for all $k \in \{\cong, \ll, \gg\}$. The converse may not be true in general. "Conjoint" local consistency, i.e., local consistency for all $k \in \{\cong, \ll, \gg\}$, does not imply consistency.[32] The implication holds under the condition in Proposition 23.

(4) What is the interplay between the local consistency of $<'_{Q|k}$ and the corresponding minimum $f_i = 0$ ($k \in \{\cong, \ll, \gg\}$, $1 \leq i \leq 3$)? It obviously holds:

$$\text{For all } (i,k) \in \{(1, \cong), (2, \ll), (3, \gg)\},\ f_i = 0 \Leftrightarrow \text{Local consistency of } <'_{Q|k}.$$

In particular, consistency implies minima $f_j(\leq,D)$=0 for all $1 \leq j \leq 3$. For f_4, minimum $f_4(\leq,D)$=0 does not hold under consistency in general. "Conjoint" minima $f_i = 0$ for all i=1,2,3 do not imply consistency in general. The latter implication holds under the condition in Proposition 23.

(5) The partial functions f_i ($1 \leq i \leq 4$) depend on trivial response patterns:

Corollary 26 (Out of proof of Lemma 24). *Let Q:=$\{I_l$: $1 \leq l \leq m\}$ ($m \geq 2$), $\leq \in \mathcal{S}$, and $D \in \mathcal{M}(n \times m;\ \{0,1\})$. For $n_0, n_1 \in \mathbb{N} \cup \{0\}$, let D^t represent D enriched with n_0 empty response patterns (0,0,. . .,0) and n_1 full response patterns (1,1,. . .,1) (i.e., as matrix rows). Then, with special bias terms F_i^*: $(\mathbb{N} \cup \{0\}) \times (\mathbb{N} \cup \{0\}) \times \mathcal{D} \times \mathcal{S} \rightarrow \mathbb{R}$ for $1 \leq i \leq 4$,*

$$f_i(\leq, D^t) = f_i(\leq, D) + F_i^*(n_0, n_1, D, \leq).$$

12.3.2. Exploratory data and structure analysis

The decomposition of CA($\leq$,D) may be utilized for descriptive data and structure analysis. We distinguish between (I) the "one-model" case (only one model analyzed), and (II) the "two-model" case (two different models compared). In both of the cases only one data matrix is fixed.[33]

[31]The definition is only formulated for $<'_{Q|\cong}$ (for $<'_{Q|\ll}$ and $<'_{Q|\gg}$ similar).

[32]The reason is that non-reflexive pairs in $\leq$ with non-existent empirical correlation are not determined by any of the elements in any of the sets $<'_{Q|k}$ ($k \in \{\cong, \ll, \gg\}$). These pairs may cause contradictions with the data (Section 10).

[33]Other "multi-model and -matrix" cases are conceivable and could be analyzed on similar grounds.

12.3.2.1. One-model case. Ad (I). Let $\leq \in \mathcal{S}$ be a model, and $D \in \mathcal{D}$ a binary data matrix.

If $f_i(\leq, D) > 0$, then the corresponding local consistency of $<'_{Q|k}$ is not fulfilled ($1 \leq i \leq 3$, $k \in \{\cong, \ll, \gg\}$):

(i=1) If $f_1(\leq, D) > 0$, then $<'_{Q|\cong}$ is not locally consistent with D. Thus, there is at least one $\leq$-equivalent[34] pair with existent r_{ij}, contradicting D. Of course, there might be $\leq$-equivalent pairs with non-existent r_{ij}, contradicting D. The existence of such pairs is not guaranteed by $f_1(\leq, D) > 0$. However, it is conceivable to have an algorithm checking for the presence of such pairs. Successively, pick the $s_i = \underline{1}_n$ and $s_i = \underline{0}_n$ columns of D, and screen the respective pairs $(J, s_i = \underline{1}_n)$ and $(s_i = \underline{0}_n, J)$ for $\leq$-equivalence and contradiction, for all other items $J \in Q$. For small structures and data sets, this could be done even by hand. If using CA($\leq$,D) within ITA, an alternative is to remove all items $I_i \in Q$ with $s_i = \underline{1}_n$ or $s_i = \underline{0}_n$ before applying ITA.

(i=2) If $f_2(\leq, D) > 0$, then there is at least one pair $(I_i, I_j) \in \leq$ which contradicts D, with (a) $i < j$, (b) existent r_{ij}, and (c) $(I_j, I_i) \notin \leq$. Of course, there might be pairs $(I_i, I_j) \in \leq$ with (a), (c), and non-existent r_{ij}, contradicting D. The existence of such pairs is not implied by $f_2(\leq, D) > 0$. Again, it is conceivable to construct an algorithm checking for the presence of such item pairs. Successively, pick the $s_i = \underline{1}_n$ and $s_i = \underline{0}_n$ columns of D, and screen all respective pairs $(I_j, s_i = \underline{1}_n)$ with $j < i$ and $(s_i = \underline{0}_n, I_j)$ with $i < j$ for "$\in$-relation" with $\leq$, (c), and contradiction. If used within ITA, an alternative is to remove all items $I_i \in Q$ with $s_i = \underline{1}_n$ or $s_i = \underline{0}_n$.

(i=3) If $f_3(\leq, D) > 0$, then there is at least one pair $(I_i, I_j) \in \leq$ which contradicts D, with (a) $j < i$, (b) existent r_{ij}, and (c) $(I_j, I_i) \notin \leq$. Of course, there might be pairs $(I_i, I_j) \in \leq$ with (a), (c), and non-existent r_{ij}, contradicting D. The existence of such pairs is not implied by $f_3(\leq, D) > 0$. One could give an algorithm checking for the presence of such pairs.

Conversely, if $f_i(\leq, D) = 0$, then the corresponding local consistency of $<'_{Q|k}$ is fulfilled ($1 \leq i \leq 3$, $k \in \{\cong, \ll, \gg\}$):

(i=1)′ If $f_1(\leq, D) = 0$, then every $\leq$-equivalent pair with existent r_{ij} does not cause contradiction. Of course, $\leq$-equivalent pairs with non-existent r_{ij} may cause contradiction. This is not determined by $f_1(\leq, D) = 0$, and can be checked separately (see (i=1)).

(i=2)′ If $f_2(\leq, D) = 0$, then none of the pairs $(I_i, I_j) \in \leq$ with (a) $i < j$, (b) existent r_{ij}, and (c) $(I_j, I_i) \notin \leq$ contradict D. There might be pairs with (a), (c), and non-existent r_{ij}, contradicting D. The existence of such pairs is not given by $f_2(\leq, D) = 0$, and can be checked separately (see (i=2)).

(i=3)′ If $f_3(\leq, D) = 0$, then there exists no pair $(I_i, I_j) \in \leq$ with (a) $j < i$, (b) existent r_{ij}, and (c) $(I_j, I_i) \notin \leq$, contradicting D. There might be pairs with (a), (c), and non-existent r_{ij}, contradicting D. This is not determined by $f_3(\leq, D) = 0$, and can be checked separately (see (i=3)).

---

[34] A pair $(I_i, I_j) \in Q \times Q$ is called $\leq$-equivalent iff $(I_i, I_j) \in \leq$ and $(I_j, I_i) \in \leq$.

12.3.2.2. Two-model case. Ad (II). Let $\leq_1,\leq_2\in\mathcal{S}$ be two models, $D\in\mathcal{D}$ a binary data matrix.

Preliminary considerations. Let $<'_Q:=\{(I_i,I_j)\in Q\times Q\colon i<j \text{ and } r_{ij} \text{ exists}\}$. Then, we have (Section 7):

$$\begin{aligned}\sum_{k\in\{\cong,\ll,\gg,\not\asymp\}} <'_{Q|_k}[\leq_1] &= <'_Q \\ &= \sum_{k\in\{\cong,\ll,\gg,\not\asymp\}} <'_{Q|_k}[\leq_2],\end{aligned}$$

with (for $s\in\{1, 2\}$)

$$\begin{aligned}
<'_{Q|_\cong}[\leq_s] &:= <'_Q \cap \{(I_i,I_j)\in Q\times Q : (I_i,I_j)\in\leq_s \wedge (I_j,I_i)\in\leq_s\},\\
<'_{Q|_\ll}[\leq_s] &:= <'_Q \cap \{(I_i,I_j)\in Q\times Q : (I_i,I_j)\in\leq_s \wedge (I_j,I_i)\notin\leq_s\},\\
<'_{Q|_\gg}[\leq_s] &:= <'_Q \cap \{(I_i,I_j)\in Q\times Q : (I_i,I_j)\notin\leq_s \wedge (I_j,I_i)\in\leq_s\},\\
<'_{Q|_{\not\asymp}}[\leq_s] &:= <'_Q \cap \{(I_i,I_j)\in Q\times Q : (I_i,I_j)\notin\leq_s \wedge (I_j,I_i)\notin\leq_s\}.
\end{aligned}$$

The families $\mathcal{F}[\leq_s]:=(<'_{Q|_\cong}[\leq_s],<'_{Q|_\ll}[\leq_s],<'_{Q|_\gg}[\leq_s],<'_{Q|_{\not\asymp}}[\leq_s])$ $(s=1,2)$ represent disjoint coverings of $<'_Q$.

Then, for all $k\in\{\cong, \ll, \gg, \not\asymp\}$ and $s,s'\in\{1,2\}$ with $s\neq s'$,

$$\begin{aligned}<'_{Q|_k}[\leq_s] &= <'_{Q|_k}[\leq_s]\cap <'_Q = <'_{Q|_k}[\leq_s]\cap \sum_{k'\in\{\cong,\ll,\gg\not\asymp\}} <'_{Q|_{k'}}[\leq_{s'}]\\ &= \sum_{k'\in\{\cong,\ll,\gg,\not\asymp\}} [<'_{Q|_k}[\leq_s]\cap <'_{Q|_{k'}}[\leq_{s'}]].\end{aligned}$$

The family $\mathcal{I}_{s,k}:=(<'_{Q|_k}[\leq_s]\cap<'_{Q|_{k'}}[\leq_{s'}]\colon k'\in\{\cong, \ll, \gg, \not\asymp\})$ is a disjoint covering of $<'_{Q|_k}[\leq_s]$. Thus, every $<'_{Q|_k}[\leq_1]$ is partitioned with respect to the four cases $<'_{Q|_{k'}}[\leq_2]$, and, respectively, every $<'_{Q|_{k'}}[\leq_2]$ with respect to the four cases $<'_{Q|_k}[\leq_1]$. Family $\mathcal{E}:=((k)_1\cap(k')_2\colon k\in\{\cong,\ll,\gg,\not\asymp\},\ k'\in\{\cong,\ll,\gg,\not\asymp\})$ is a *finer* disjoint covering of $<'_Q$ than $\mathcal{F}[\leq_1]$ and $\mathcal{F}[\leq_2]$ (See Table 1).[35]

For $\leq_s$ $(s=1,2)$,

$$\mathrm{CA}(\leq_s,D) = 1-\frac{2}{m(m-1)}\sum_{k=1}^{4} f_k(\leq_s,D),$$

[35] $<'_{Q|_\cong}[\leq_s],<'_{Q|_\ll}[\leq_s]$, etc. $(s=1, 2)$ are also written $\cong_s$, $\ll_s$, etc. Note that "Total" is meant in the sense of set-theoretic union (here, even Σ).

Table 1
Partitioning sets

$\leq_1 \backslash \leq_2$	$<'_{Q\|\cong}[\leq_2]$	$<'_{Q\|\ll}[\leq_2]$	$<'_{Q\|\gg}[\leq_2]$	$<'_{Q\|\not\asymp}[\leq_2]$	Total (Σ)
$<'_{Q\|\cong}[\leq_1]$	$\cong_1 \cap \cong_2$	$\cong_1 \cap \ll_2$	$\cong_1 \cap \gg_2$	$\cong_1 \cap \not\asymp_2$	$<'_{Q\|\cong}[\leq_1]$
$<'_{Q\|\ll}[\leq_1]$	$\ll_1 \cap \cong_2$	$\ll_1 \cap \ll_2$	$\ll_1 \cap \gg_2$	$\ll_1 \cap \not\asymp_2$	$<'_{Q\|\ll}[\leq_1]$
$<'_{Q\|\gg}[\leq_1]$	$\gg_1 \cap \cong_2$	$\gg_1 \cap \ll_2$	$\gg_1 \cap \gg_2$	$\gg_1 \cap \not\asymp_2$	$<'_{Q\|\gg}[\leq_1]$
$<'_{Q\|\not\asymp}[\leq_1]$	$\not\asymp_1 \cap \cong_2$	$\not\asymp_1 \cap \ll_2$	$\not\asymp_1 \cap \gg_2$	$\not\asymp_1 \cap \not\asymp_2$	$<'_{Q\|\not\asymp}[\leq_1]$
Total (Σ)	$<'_{Q\|\cong}[\leq_2]$	$<'_{Q\|\ll}[\leq_2]$	$<'_{Q\|\gg}[\leq_2]$	$<'_{Q\|\not\asymp}[\leq_2]$	$<'_Q$

with (Z:=$[a_{ij}+b_{ij}]\cdot[c_{ij}+d_{ij}]\cdot[a_{ij}+c_{ij}]\cdot[b_{ij}+d_{ij}]$)

$$f_1(\leq_s, D) := \sum_{(I_i,I_j)\in <'_{Q|\cong}[\leq_s]} \left[\frac{a_{ij}d_{ij} - b_{ij}c_{ij}}{\sqrt{Z}} - 1\right]^2,$$

$$f_2(\leq_s, D) := \sum_{(I_i,I_j)\in <'_{Q|\ll}[\leq_s]} \frac{[nc_{ij}]^2}{Z},$$

$$f_3(\leq_s, D) := \sum_{(I_i,I_j)\in <'_{Q|\gg}[\leq_s]} \frac{[nb_{ij}]^2}{Z},$$

$$f_4(\leq_s, D) := \sum_{(I_i,I_j)\in <'_{Q|\not\asymp}[\leq_s]} \frac{[a_{ij}d_{ij} - b_{ij}c_{ij}]^2}{Z}.$$

Next, we define a table of local summary statistics corresponding to Table 1. Consider ($\cong_1 \cap \ll_2$). For a pair (I_i,I_j) in ($\cong_1 \cap \ll_2$), $[I_i \leq_1 I_j \wedge I_j \leq_1 I_i]$ and $[I_i \leq_2 I_j \wedge I_j \not\leq_2 I_i]$. There are two prescriptions for calculating theoretical correlation r^*_{ij}: (a) viewed in regard to $\leq_1$, and (b) with respect to $\leq_2$. In case of (a), $\delta^2_{ij} := (r_{ij} - r^*_{ij})^2$ is $\left[\frac{a_{ij}d_{ij}-b_{ij}c_{ij}}{\sqrt{Z}} - 1\right]^2$, whereas for (b), it is $[nc_{ij}]^2/Z$. Summation over all pairs in ($\cong_1 \cap \ll_2$) yields $\sum_{(I_i,I_j)\in(\cong_1\cap\ll_2)} \left[\frac{a_{ij}d_{ij}-b_{ij}c_{ij}}{\sqrt{Z}} - 1\right]^2$ in case of (a), and, respectively, $\sum_{(I_i,I_j)\in(\cong_1\cap\ll_2)} \frac{[nc_{ij}]^2}{Z}$ in case of (b). Let these two functions be denoted by $f_{\cong_{\hat{1}},\ll_2}$ [case (a)], and respectively $f_{\cong_1,\ll_{\hat{2}}}$ [case (b)].[36] For the sets on the diagonal of Table 1, i.e., ($\cong_1 \cap \cong_2$), ($\ll_1 \cap \ll_2$), ($\gg_1 \cap \gg_2$), and ($\not\asymp_1 \cap \not\asymp_2$), both prescriptions coincide. Thus, only one measure is determined for any of these sets. Let them be denoted by $f_{\cong 1, \cong 2}$, $f_{\ll 1, \ll 2}$, etc. See Table 2.[37]

Issues like boundedness, consistency–maximum problem, etc., already discussed for global measure CA($\leq$,D) and 1st-order local measures f_i, can be transfered to and re-discussed for the finer 2nd-order local measures $f_{\cong_1,\cong_2}$, $f_{\cong_{\hat{1}},\ll_2}$, $f_{\cong_1,\ll_{\hat{2}}}$, etc. (partly, with minor modifications).[38]

[36]The symbol ^ indicates the model the calculation is taken in regard to.

[37]Row and column totals are obtained by summing over $\leq_1$- and $\leq_2$-calculated f values, respectively (e.g., $f_1(\leq_1,D) = f_{\cong_1,\cong_2} + f_{\cong_{\hat{1}},\ll_2} + f_{\cong_{\hat{1}},\gg_2} + f_{\cong_{\hat{1}},\not\asymp_2}$).

[38]Note that for "non-diagonal" sets in Table 1, local consistency can be defined in regard to model $\leq_1$, and with respect to model $\leq_2$.

Table 2
Local statistics

$\leq_1 \backslash \leq_2$	$<'_{Q\mid_{\cong}} [\leq_2]$	$<'_{Q\mid_{\ll}} [\leq_2]$	$<'_{Q\mid_{\gg}} [\leq_2]$	$<'_{Q\mid_{\not\cong}} [\leq_2]$	Total (Σ)
$<'_{Q\mid_{\cong}} [\leq_1]$	$f_{\cong_1,\cong_2}$	$f_{\cong_{\hat 1},\ll_2} \mid f_{\cong_1,\ll_{\hat 2}}$	$f_{\cong_{\hat 1},\gg_2} \mid f_{\cong_1,\gg_{\hat 2}}$	$f_{\cong_{\hat 1},\not\cong_2} \mid f_{\cong_1,\not\cong_{\hat 2}}$	$f_1(\leq_1,D)$
$<'_{Q\mid_{\ll}} [\leq_1]$	$f_{\ll_{\hat 1},\cong_2} \mid f_{\ll_1,\cong_{\hat 2}}$	$f_{\ll_1,\ll_2}$	$f_{\ll_{\hat 1},\gg_2} \mid f_{\ll_1,\gg_{\hat 2}}$	$f_{\ll_{\hat 1},\not\cong_2} \mid f_{\ll_1,\not\cong_{\hat 2}}$	$f_2(\leq_1,D)$
$<'_{Q\mid_{\gg}} [\leq_1]$	$f_{\gg_{\hat 1},\cong_2} \mid f_{\gg_1,\cong_{\hat 2}}$	$f_{\gg_{\hat 1},\ll_2} \mid f_{\gg_1,\ll_{\hat 2}}$	$f_{\gg_1,\gg_2}$	$f_{\gg_{\hat 1},\not\cong_2} \mid f_{\gg_1,\not\cong_{\hat 2}}$	$f_3(\leq_1,D)$
$<'_{Q\mid_{\not\cong}} [\leq_1]$	$f_{\not\cong_{\hat 1},\cong_2} \mid f_{\not\cong_1,\cong_{\hat 2}}$	$f_{\not\cong_{\hat 1},\ll_2} \mid f_{\not\cong_1,\ll_{\hat 2}}$	$f_{\not\cong_{\hat 1},\gg_2} \mid f_{\not\cong_1,\gg_{\hat 2}}$	$f_{\not\cong_1,\not\cong_2}$	$f_4(\leq_1,D)$
Total (Σ)	$f_1(\leq_2,D)$	$f_2(\leq_2,D)$	$f_3(\leq_2,D)$	$f_4(\leq_2,D)$	–

Actual elaborations. Each of the models $\leq_1$ and $\leq_2$ could be investigated, conditional on the other model given. This corresponds to a "one-model analysis" as sketched in (I), however, this time, based on the finer 2nd-order local measures $f_{\cong_1,\cong_2}$, $f_{\cong_{\hat 1},\ll_2}$, $f_{\cong_1,\ll_{\hat 2}}$, etc. This is straightforward, and is left to the reader.

Models $\leq_1$ and $\leq_2$ can be compared, *locally*, in regard to their *descriptive adequacy* with data D. This is sketched next, using Table 2:[39]

First row $\cong_1$. *Column* $\ll_2$: (1) Let $f_{\cong_{\hat 1},\ll_2}=0$. Then, $f_{\cong_1,\ll_{\hat 2}}=0$. Consider a pair $(I_i,I_j)\in(\cong_1\cap\ll_2)$. Then, (I_i,I_j) and (I_j,I_i) are in $\leq_1$, whereas only (I_i,I_j) but not (I_j,I_i) belongs to $\leq_2$. $\leq_2$. Since $f_{\cong_{\hat 1},\ll_2}=0$, neither (I_i,I_j) nor (I_j,I_i) contradicts D. The pair (I_j,I_i) might be a candidate for $\leq_2$.[40] (2) Let $f_{\cong_{\hat 1},\ll_2}>0$. If $f_{\cong_1,\ll_{\hat 2}}=0$, then pair (I_i,I_j) in $(\cong_1\cap\ll_2)$ should rather be classified according to $\leq_2$- than $\leq_1$-prescription:[41] For every pair $(I_i,I_j)\in(\cong_1\cap\ll_2)$, the "direction" (I_i,I_j) does not cause contradiction. However, for an appropriate $(I_i,I_j)\in(\cong_1\cap\ll_2)$, there is a pair (I_j,I_i) contradicting D. These "directions" $(I_j,I_i)\in\leq_1$ (for $(I_i,I_j)\in(\cong_1\cap\ll_2)$) might be candidates to exclude from $\leq_1$.

Column $\gg_2$: This is dual to (1) and (2), Column $\ll_2$.

Column $\not\cong_2$: Let $f_{\cong_{\hat 1},\not\cong_2}=0$. For any $(I_i,I_j)\in(\cong_1\cap\not\cong_2)$, neither (I_i,I_j) $(\notin\leq_2)$ nor (I_j,I_i) $(\notin\leq_2)$ contradicts D. These pairs might be candidates for $\leq_2$.

Second row $\ll_1$. *Column* $\cong_2$: This is dual to First row $\cong_1$, Column $\ll_2$.

Column $\gg_2$: (1) Let $f_{\ll_{\hat 1},\gg_2}>0$. If $f_{\ll_1,\gg_{\hat 2}}=0$, then pair (I_i,I_j) in $(\ll_1\cap\gg_2)$ should be classified according to $\leq_1$- than $\leq_2$-prescription:[42] For every pair $(I_i,I_j)\in(\ll_1\cap\gg_2)$, the "direction" (I_i,I_j) does not cause contradiction. However, for an appropriate $(I_i,I_j)\in(\ll_1\cap\gg_2)$, there is a pair (I_j,I_i) which contradicts D. These "directions" $(I_j,I_i)\in\leq_2$ and $(I_i,I_j)\notin\leq_2$ (for $(I_i,I_j)\in(\ll_1\cap\gg_2)$) might be candidates to respectively exclude from or include into $\leq_2$. (2) Let $f_{\ll_{\hat 1},\gg_2}=0$. If $f_{\ll_1,\gg_{\hat 2}}=0$, for every pair $(I_i,I_j)\in(\ll_1\cap\gg_2)$, the "directions" $(I_i,I_j)\notin\leq_2$ and $(I_j,I_i)\notin\leq_1$ might be candidates to include into $\leq_2$ and $\leq_1$, respectively. (3) Let $f_{\ll_{\hat 1},\gg_2}>0$. If $f_{\ll_1,\gg_{\hat 2}}=0$, then pair (I_i,I_j) in $(\ll_1\cap\gg_2)$ should be classified according to $\leq_2$- than $\leq_1$-prescription.

[39] We only focus on some "important" issues.

[40] For example, in the context of a data-analytic search method for surmise relations.

[41] That is, $[(I_i,I_j)\in\leq_2\wedge(I_j,I_i)\notin\leq_2]$- than $[(I_i,I_j)\in\leq_1\wedge(I_j,I_i)\in\leq_1]$-prescription.

[42] That is, $[(I_i,I_j)\in\leq_1\wedge(I_j,I_i)\notin\leq_1]$- than $[(I_i,I_j)\notin\leq_2\wedge(I_j,I_i)\in\leq_2]$-prescription.

Column $\not\asymp_2$: Let $f_{\ll_1,\not\asymp_2}=0$. For any $(I_i,I_j)\in(\ll_1\cap\not\asymp_2)$, (I_i,I_j) $(\notin\leq_2)$ does not contradict D. These pairs might be candidates for $\leq_2$.

Third and fourth rows $\gg_1$ *and* $\not\asymp_1$. This is dual to the previous cases.

The descriptive arguments in Section 12.3 may lead to revisions in a priori hierarchical dependencies between bi-valued test items. This might play a role in *search* and *discovery algorithms* for detecting relations among test items. Even more, the arguments could possibly be used to elaborate an alternative data-analytic method.

12.4. Final general remarks

The coefficient CA($\leq$,D) is not derived based on a clear statistical model. It is descriptive, no sampling and inference are considered. The measure lacks a clear population analogue; no reasonable point estimators with (asymptotic) distributions, no confidence intervals, no hypotheses testing, and no statistical comparisons between different population values are provided.

The coefficient CA($\leq$,D) does not satisfy certain methodological issues desirable for association measures in general: (1) A certain value of CA($\leq$,D) is not interpretable sensibly in terms of concepts of a statistical model. (2) The measure does not sufficiently take into account the nominal character of the data. The Pearson correlation (respectively Phi coefficient) is less than adequate (Pearson correlation, actually designed for real variables).

The surmise relation optimal with respect to CA($\leq$,D) might be used for a special (operational) purpose, e.g., adaptive assessment of knowledge. Then, the fit criterion underlying CA($\leq$,D) (i.e., agreement between r_{ij} and r^*_{ij}) may not reflect the intended purpose(s) adequately.

Future work could involve interpretable goodness-of-fit measures in regard to purpose-related fit criteria derived on the basis of the ‘‘Proportional Reduction in Predictive Error’’ approach.[43] These measures could take into account the niveau of the data. Another direction of research is to combine goodness-of-fit measures to different fit criteria to create a single measure, providing a ‘‘trade-off’’ (in whatever sense) between the fit criteria meshed. The question is how to base this on statistical grounds.[44] This would be of interest (e.g.) in KST (see Section 1.1).

Acknowledgements

This research was supported by grants from the University of Graz to Ali Ünlü.

[43] This approach was proposed originally by Guttman (1941), and was used in the series of papers by Goodman and Kruskal (1954, 1959, 1963, 1972).

[44] Existing statistical tests for the validation and comparison of probabilistic models for hierarchies among bi-valued items are based on chi-square and log-likelihood ratio statistics (see Dayton and Macready, 1976; Doignon and Falmagne, 1999). These tests do not account for purpose-specific fit criteria and ‘‘trade-off’’ among the latter. Further, they depend on sufficiently large data sets, whereas carefully performed experiments require small data sets.

Appendix A. Proofs

A.1. Proofs in Section 4

Lemma 9. Let $(I_i,I_j)\in\leq$ $(1\leq i, j\leq m)$ with $\mathbf{A}$.[45]
Consider the corresponding 2×2 table:

$\mathbf{I}_i\diagdown\mathbf{I}_j$	**1**	**0**
1	a	b
0	c	d

with entries $a,b,c,d\in\mathbb{N}\cup\{0\}$,

$$\begin{aligned}
a &= |\{k\in\{1,2,\ldots,n\} : d'_{ki}=1\wedge d'_{kj}=1\}|,\\
b &= |\{k\in\{1,2,\ldots,n\} : d'_{ki}=1\wedge d'_{kj}=0\}|,\\
c &= |\{k\in\{1,2,\ldots,n\} : d'_{ki}=0\wedge d'_{kj}=1\}|,\\
d &= |\{k\in\{1,2,\ldots,n\} : d'_{ki}=0\wedge d'_{kj}=0\}|.
\end{aligned}$$

According to Lemma 7, we have

$$r_{ij}=\frac{ad-bc}{\sqrt{(a+c)\cdot(b+d)\cdot(a+b)\cdot(c+d)}}.$$

Case 1 ($[(I_j,I_i)\in\leq]$). Since $\leq$ is consistent with D, we have $b=c=0$. Thus,

$$\begin{aligned}
r_{ij}&=\frac{ad-bc}{\sqrt{(a+c)\cdot(b+d)\cdot(a+b)\cdot(c+d)}}\\
&=\frac{ad-0\cdot 0}{\sqrt{(a+0)\cdot(0+d)\cdot(a+0)\cdot(0+d)}}=1.
\end{aligned}$$

[45]The symbol $\mathbf{A}$ is introduced in Definition 5.

Case 2 ($[(I_j,I_i)\notin\leq]$). Since $\leq$ is a total fit to D, we have $c=0$. Thus,

$$\begin{aligned} r_{ij} &= \frac{ad-bc}{\sqrt{(a+c)\cdot(b+d)\cdot(a+b)\cdot(c+d)}} \\ &= \frac{ad-b\cdot 0}{\sqrt{(a+0)\cdot(b+d)\cdot(a+b)\cdot(0+d)}} = \frac{ad}{\sqrt{ad\cdot(b+d)\cdot(a+b)}} \\ &= \sqrt{\frac{ad}{(b+d)\cdot(a+b)}} \overset{\text{(i)}}{=} \sqrt{\frac{(1-p_{I_i})\cdot p_{I_j}}{(1-p_{I_j})\cdot p_{I_i}}}. \end{aligned}$$

Ad 27 (i). We see that by direct transformations: $\frac{ad}{(b+d)\cdot(a+b)} = \frac{(1-p_{I_i})\cdot p_{I_j}}{(1-p_{I_j})\cdot p_{I_i}}$. □

Lemma 11. This example is proposed originally by Schrepp (1999, pp. 364–365): Let $Q:=\{I_1,I_2,I_3,I_4\}$ be a set of four dichotomous items. The binary relation $\leq$ is a surmise relation on Q,[46]

$$\leq:=\Delta\cup\{(I_2,I_1),(I_3,I_1),(I_4,I_1),(I_4,I_3)\}.$$

Consider the data matrix $D\in\mathcal{M}(7\times 4;\ \{0,1\})$,

$$D:=\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

Then, $\leq$ is consistent with D. The pair $(I_2,I_4)\in Q\times Q$ is not-$\leq$-comparable satisfying **A**. For r_{24}, $r_{24}=[\sqrt{120}]^{-1}$ (≈ 0.09). Thus, $r_{24}\neq 0$ $(=:r^*_{24})$. □

Proposition 12. Let $[(I_i,I_j)\notin\leq\wedge(I_j,I_i)\in\leq]$.[47] By means of the definitions of empirical and theoretical correlation, both concepts of correlation are "symmetric in the indexes", i.e., for all $(I_i,I_j)\in[Q\times Q]_{\mathbf{A}}$, $r_{ij}=r_{ji}$ and $r^*_{ij}=r^*_{ji}$. Therefore,

$$\delta_{ij}:=r_{ij}-r^*_{ij}=r_{ji}-r^*_{ji}\overset{\text{(i)}}{=}0.$$

Ad (i). Since $(I_j,I_i)\in\leq$ with **A**, Corollary 10 assures $r_{ji}-r^*_{ji}=0$. □

[46] Δ denotes the diagonal $\Delta:=\{J=(J_1,J_2)\in Q\times Q\colon J_1=J_2\}$ in $Q\times Q$.

[47] If $[(I_i,I_j)\in\leq\vee(I_j,I_i)\notin\leq]$, it results from Corollary 10 and Lemma 11.

A.2. Proofs in Section 5

Proposition 13 (**RIN**). Of course, $0 \leq r_{ij}^*$. Show: $r_{ij}^* \leq n-1$.
Consider cases $[(I_i,I_j) \in \leq \wedge (I_j,I_i) \notin \leq]$ and $[(I_i,I_j) \notin \leq \wedge (I_j,I_i) \in \leq]$:[48]

Case 1 $[(I_i,I_j) \in \leq \wedge (I_j,I_i) \notin \leq]$. Then, we have

$$r_{ij}^* := \sqrt{\frac{(1-p_{I_j}) \cdot p_{I_j}}{(1-p_{I_i}) \cdot p_{I_i}}} \overset{(i)}{\leq} \sqrt{\frac{(1-n^{-1}) \cdot n^{-1}(n-1)}{[1-n^{-1}(n-1)] \cdot n^{-1}}} = \sqrt{(n-1)^2} = n-1.$$

Ad (i). Since $(I_i,I_j) \in [Q \times Q]_{\mathbf{A}}$, it holds $n^{-1} \leq p_{I_l} \leq n^{-1}(n-1)$, $l=i,j$.

Case 2 $[(I_i,I_j) \notin \leq \wedge (I_j,I_i) \in \leq]$. Then, we have

$$r_{ij}^* := \sqrt{\frac{(1-p_{I_i}) \cdot p_{I_i}}{(1-p_{I_j}) \cdot p_{I_j}}} \leq \sqrt{\frac{(1-n^{-1}) \cdot n^{-1}(n-1)}{[1-n^{-1}(n-1)] \cdot n^{-1}}} = \sqrt{(n-1)^2} = n-1.$$

(**PDt+∞**). Let $m \in \mathbb{N}$, $m \geq 2$. Let $I_1,I_2 \in Q$. Consider the surmise relation $\leq_* := \Delta \cup \{(I_1,I_2)\}$ ($\Delta := \{J=(J_1,J_2) \in Q \times Q: J_1=J_2\}$).

We show that this special choice of $\leq_*$ and (I_1,I_2) with $1<2$, $(I_1,I_2) \notin \leq_*$, and $(I_2,I_1) \in \leq_*$ fulfills the required condition

$$\forall n \geq 2 \exists D_{n-1} \in \mathcal{M}(n \times m; \{0,1\}) : [(r_{12})_{n-1} = -1] \wedge [(r_{12}^*)_{n-1} = n-1].$$

For every $n \in \mathbb{N}$, $n \geq 2$, define the $n \times m$ matrix $D_{n-1} \in \mathcal{M}$ $(n \times m;\ \{0,1\})$,[49]

$$D_{n-1} := \begin{pmatrix} 1 & 0 & * & \cdots & * & * & * \\ 0 & 1 & * & \ddots & & & \vdots \\ 0 & 1 & * & & \ddots & & * \\ \vdots & \vdots & \vdots & & & \ddots & * \\ 0 & 1 & * & * & * & \cdots & * \end{pmatrix}.$$

The column s_1 of D_{n-1} corresponding to $I_1 \in Q$ is

$$s_1 := (1, \underbrace{0, 0, \ldots, 0}_{(n-1) times})^{\mathrm{T}},$$

[48]If $[(I_i,I_j) \in \leq \wedge (I_j,I_i) \in \leq]$, then, obviously, $r_{ij}^* := 1 \leq n-1$ (note: $n \geq 2$). If $[(I_i,I_j) \notin \leq \wedge (I_j,I_i) \notin \leq]$, then $r_{ij}^* := 0 \leq n-1$.
[49]The symbol $*$ substitutes 0 and 1, arbitrarily combined.

and the column s_2 of D_{n-1} corresponding to $I_2 \in Q$ is

$$s_2 := (0, \underbrace{1, 1, \ldots, 1}_{(n-1) times})^{\mathrm{T}}.$$

Let $n \in \mathbb{N}$, $n \geq 2$, and $D_{n-1} \in \mathcal{M}$ ($n \times m$; $\{0,1\}$) (previously defined). The empirical correlation $(r_{12})_{n-1}$ between $I_1, I_2 \in Q$ under D_{n-1} satisfies $(r_{12})_{n-1} = -1$. The 2×2 table for (I_1, I_2) is given by

$\mathbf{I_1 \diagdown I_2}$	**1**	**0**
1	0	1
0	$n-1$	0

According to Lemma 7, we have

$$(r_{12})_{n-1} = \frac{0 \cdot 0 - (n-1) \cdot 1}{\sqrt{(n-1) \cdot 1 \cdot 1 \cdot (n-1)}} = -1.$$

The theoretical correlation $(r_{12}^*)_{n-1}$ between $I_1, I_2 \in Q$ under D_{n-1} reduces to $(r_{12}^*)_{n-1} = n-1$. Since $(I_1, I_2) \in \leq_*$ and $(I_2, I_1) \notin \leq_*$, we have

$$(r_{12}^*)_{n-1} := \sqrt{\frac{(1 - p_{I_1}) \cdot p_{I_2}}{(1 - p_{I_2}) \cdot p_{I_1}}}.$$

We see that $p_{I_1} = n^{-1}$ and $p_{I_2} = n^{-1}(n-1)$. Finally, it holds

$$(r_{12}^*)_{n-1} = \sqrt{\frac{(1 - n^{-1}) \cdot n^{-1}(n-1)}{[1 - n^{-1}(n-1)] \cdot n^{-1}}} = n - 1.$$

The corresponding sequence $((r_{12}^*)_n)_{n \in \mathbb{N}}$, $(r_{12}^*)_n = n$ for all $n \in \mathbb{N}$, diverges properly to $+\infty$, i.e., $\lim_{n \to \infty} ((r_{12}^*)_n)_{n \in \mathbb{N}} = +\infty$. □

Corollary 14. Since $r_{ij} \in [-1,1]$ and $r_{ij}^* \in [0, n-1]$, we have:

$$\delta_{ij} := r_{ij} - r_{ij}^* \leq 1 - r_{ij}^* \leq 1 - 0 = 1,$$

and

$$\delta_{ij} := r_{ij} - r_{ij}^* \geq -1 - r_{ij}^* \geq -1 - (n-1) = -n.$$

In particular, we have $\delta_{ij}^2 \in [0, n^2]$. □

A.3. Proofs in Section 7

Lemma 17. Use $<'_Q = \sum_{k\in\{\cong,\ll,\gg,\not\asymp\}} <'_{Q|_k}$ to obtain:

$$
\begin{aligned}
1-\mathrm{CA}(\leq,D) &= \frac{2}{m(m-1)}\sum_{(I_i,I_j)\in<'_Q}(r_{ij}-r^*_{ij})^2 = \frac{2}{m(m-1)}\sum_{(I_i,I_j)\in<'_{Q|\cong}}(r_{ij}-r^*_{ij})^2 \\
&\quad+\frac{2}{m(m-1)}\sum_{(I_i,I_j)\in<'_{Q|\ll}}(r_{ij}-r^*_{ij})^2+\frac{2}{m(m-1)}\sum_{(I_i,I_j)\in<'_{Q|\gg}}(r_{ij}-r^*_{ij})^2 \\
&\quad+\frac{2}{m(m-1)}\sum_{(I_i,I_j)\in<'_{Q|\not\asymp}}(r_{ij}-r^*_{ij})^2.
\end{aligned}
$$

Substitute $\frac{a_{ij}d_{ij}-b_{ij}c_{ij}}{\sqrt{(a_{ij}+c_{ij})\cdot(b_{ij}+d_{ij})\cdot(a_{ij}+b_{ij})\cdot(c_{ij}+d_{ij})}}$ for r_{ij} to obtain $f_1(\leq,D)$, $f_4(\leq,D)$ trivially, and $f_2(\leq,D)$, $f_3(\leq,D)$ after summarizing appropriate terms. □

A.4. Proofs in Section 8

Proposition 18 (**RIN**). Let $\leq\in$S. Let $<'_Q\neq\emptyset$. Then,

$$
\mathrm{CA}(\leq,D):=1-\frac{2}{m(m-1)}\sum_{(I_i,I_j)\in<'_Q}(r_{ij}-r^*_{ij})^2\overset{\text{(i)}}{\leq}1.
$$

Ad (i). Clearly, $2[m(m-1)]^{-1}\sum_{(I_i,I_j)\in<'_Q}(r_{ij}-r^*_{ij})^2\geq 0$.

$$
\begin{aligned}
\mathrm{CA}(\leq,D)&:=1-\frac{2}{m(m-1)}\sum_{(I_i,I_j)\in<'_Q}(r_{ij}-r^*_{ij})^2\overset{\text{(i)}}{\geq}1-\frac{2}{m(m-1)}\sum_{(I_i,I_j)\in<'_Q}n^2 \\
&=1-\frac{2}{m(m-1)}\cdot|<'_Q|\cdot n^2\overset{\text{(ii)}}{\geq}1-\frac{2}{m(m-1)}\cdot\frac{m(m-1)}{2}\cdot n^2=1-n^2.
\end{aligned}
$$

Ad (i). According to Corollary 14, $0\leq(r_{ij}-r^*_{ij})^2\leq n^2$ for all $(I_i,I_j)\in<'_Q$.

Ad (ii). We have $|<'_Q|\leq 2^{-1}[m(m-1)]$.

(PDt$-\infty$). According to Proposition 13, there is a surmise relation $\leq_*$ and a sequence $(D_n)_{n\in\mathbb{N}}$ with: For an appropriate $(I_{i_0},I_{j_0})\in Q\times Q$ with $i_0<j_0$, $[(I_{i_0},I_{j_0})\in\leq_*\wedge(I_{j_0},I_{i_0})\notin\leq_*]$, and $[\forall n\in\mathbb{N}$: (I_{i_0},I_{j_0}) fulfills **A** under $D_n]$, it holds $((r_{i_0j_0})_n)_{n\in\mathbb{N}}=(-1,-1,-1,-1,\ldots)$ and $((r^*_{i_0j_0})_n)_{n\in\mathbb{N}}=(1,2,3,4,\ldots)$.[50]

[50]For $n\in\mathbb{N}$, $(r_{i_0j_0})_n$ and $(r^*_{i_0j_0})_n$ respectively denote empirical and theoretical correlation between I_{i_0} and I_{j_0} under D_n.

These facts imply ($n\in\mathbb{N}$ arbitrary):[51]

$$\begin{aligned}\mathrm{CA}(\leq_*, D_n) &:= 1 - \frac{2}{m(m-1)} \sum_{(I_i,I_j)\in <'_Q(D_n)} (r_{ij} - r^*_{ij})^2 \\ &\overset{\text{(i)}}{\leq} 1 - \frac{2}{m(m-1)} ((r_{i_0 j_0})_n - (r^*_{i_0 j_0})_n)^2 \\ &\overset{\text{(ii)}}{=} 1 - \frac{2(n+1)^2}{m(m-1)}.\end{aligned}$$

Ad (i). $(I_{i_0}, I_{j_0})\in <'_Q(D_n)$: $((r_{i_0 j_0})_n - (r^*_{i_0 j_0})_n)^2 \leq \sum_{(I_i,I_j)\in <'_Q(D_n)} (r_{ij} - r^*_{ij})^2$.
Ad (ii). It is the case that $(r_{i_0 j_0})_n = -1$ and $(r^*_{i_0 j_0})_n = n$.
The sequence $(1-[2(n+1)^2]/[m(m-1)])_{n\in\mathbb{N}}$ diverges properly to $-\infty$. □

A.5. Proofs in Section 9

Lemma 20. See the example in Lemma 11: $(r_{24}-r^*_{24})^2>0$. □

A.6. Proofs in Section 10

Lemma 22. Consider the counterexample:

$$\begin{aligned}&Q := \{I_1, I_2, I_3\},\\ &\leq := \Delta \cup \{(I_1,I_2),(I_3,I_2)\},\\ &D := \begin{bmatrix} 1 & 1 & 0 \\ \\ 0 & 0 & 0 \end{bmatrix}_{2\times 3}.\end{aligned}$$

Then, $\leq$ is not consistent with D.[52] Since $<'_Q = \{(I_1,I_2)\}$ and $r_{12}=r^*_{12}$ (=1), it holds $\mathrm{CA}(\leq,D)=1-1/3(r_{12}-r^*_{12})^2=1$. □

Proposition 23. Let $(I_i,I_j)\in\leq$ be a non-reflexive pair for which r_{ij} exists (i.e., with **A**). Show that $|\{k\in\{1,2,\ldots,n\}: d'_{ki}=0 \wedge d'_{kj}=1\}|=0$. According to Lemma 17, $\mathrm{CA}(\leq,D)=1$ is equivalent to the conjunction of the four conditions ($Z:=[a_{ij}+b_{ij}]\cdot[c_{ij}+d_{ij}]\cdot[a_{ij}+c_{ij}]\cdot[b_{ij}+d_{ij}]$):

$$\begin{aligned}&\forall (I_i,I_j)\in <'_{Q|\cong} : \frac{a_{ij}d_{ij} - b_{ij}c_{ij}}{\sqrt{Z}} - 1 = 0,\\ &\forall (I_i,I_j)\in <'_{Q|\ll} : nc_{ij} = 0,\\ &\forall (I_i,I_j)\in <'_{Q|\gg} : nb_{ij} = 0,\\ &\forall (I_i,I_j)\in <'_{Q|\not\cong} : a_{ij}d_{ij} - b_{ij}c_{ij} = 0.\end{aligned}$$

[51] We have $<'_Q(D_n) := \{(I_i,I_j)\in Q\times Q: i<j \text{ and } (I_i,I_j) \text{ fulfills } \mathbf{A} \text{ under } D_n\}$.
[52] This is only due to the pair $(I_3,I_2)\in\leq$, for which r_{32} does not exist.

Case 1 $[i<j\wedge(I_j,I_i)\notin\leq]$. Then, $(I_i,I_j)\in<'_{Q|\ll}$. Thus, $c_{ij}=0$.

Case 2 $[i<j\wedge(I_j,I_i)\in\leq]$. In this case, $(I_i,I_j)\in<'_{Q|\cong}$. The pair (I_i,I_j) is subjected to condition $\frac{a_{ij}d_{ij}-b_{ij}c_{ij}}{\sqrt{Z}}-1=0$. In other words (Lemma 7),

$$\left|\sum_{k=1}^{n}[((s_i)_k-\overline{s_i})((s_j)_k-\overline{s_j})]\right| = \sqrt{\sum_{k=1}^{n}[(s_i)_k-\overline{s_i}]^2}\cdot\sqrt{\sum_{k=1}^{n}[(s_j)_k-\overline{s_j}]^2}.$$

We apply the Cauchy–Schwarz Inequality (CSI); see Fischer (1995).

Lemma 27 [Cauchy–Schwarz Inequality]. *Let V be an inner product space with corresponding scalar product $\langle .,.\rangle$: $V\times V\to\mathbb{K}$. Then, for all $v,w\in V$,*

$$|\langle v,w\rangle|\leq\|v\|\cdot\|w\|,$$

whereupon $\|v\|:=\sqrt{\langle v,v\rangle}$ and $\|w\|:=\sqrt{\langle w,w\rangle}$ denote the norms of v and w under the canonical norm $\|.\|:V\to\mathbb{R},x\mapsto\|x\|:=\sqrt{\langle x,x\rangle}$ on V. Equality holds if and only if the vectors v and w are linearly dependent, i.e., one is a scalar multiple of the other.

Let $\langle .,.\rangle$: $\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$, $(x,y)\mapsto\langle x,y\rangle:=\sum_{k=1}^{n}x_ky_k$ be the canonical scalar product on $\mathbb{R}^n$. The corresponding canonical norm is given by

$$\|.\|:\mathbb{R}^n\to\mathbb{R},x\mapsto\|x\|:=\sqrt{\langle x,x\rangle}.$$

In this notation, the equation preceding Lemma 27 is rewritten[53]

$$|\langle s_i-\underline{(\overline{s_i})},s_j-\underline{(\overline{s_j})}\rangle| = \|s_i-\underline{(\overline{s_i})}\|\cdot\|s_j-\underline{(\overline{s_j})}\|.$$

According to the CSI, for an appropriate $\lambda\in\mathbb{R}$, one of the vectors $s_i-\underline{(\overline{s_i})}$ and $s_j-\underline{(\overline{s_j})}$ is a λ-scalar multiple of the other. Since (I_i,I_j) satisfies **A**, it holds $\lambda\neq0$. Thus, $s_i-\underline{(\overline{s_i})}=\mu\cdot[s_j-\underline{(\overline{s_j})}]$ for an appropriate $\mu\in\mathbb{R}$, $\mu\neq0$.

Step 1. There is an index $k\in\{1,2,\ldots,n\}$ with $(s_i)_k=(s_j)_k=0$.

Assumption. There is no index $k\in\{1,2,\ldots,n\}$ with $(s_i)_k=(s_j)_k=0$. The pair (I_i,I_j) satisfies **A**. Therefore, there are $k_1,k_2\in\{1,2,\ldots,n\}$ with $(s_i)_{k_1}=0$ and $(s_i)_{k_2}=1$. Under assumption, $(s_j)k_1=1$.

Case 1 $[(s_j)_{k_2}=0]$. Then, $0=\mu-\mu\overline{s_j}+\overline{s_i}$ and $1=\overline{s_i}-\mu\overline{s_j}$. Thus, $\mu=-1$. This implies $s_i-\underline{(\overline{s_i})}=(-1)\cdot[s_j-\underline{(\overline{s_j})}]$. Then, we obtain a contradiction:

$$\begin{aligned}0&\overset{\text{(i)}}{<}\langle s_i-\underline{(\overline{s_i})},s_j-\underline{(\overline{s_j})}\rangle\\&=\langle(-1)\cdot[s_j-\underline{(\overline{s_j})}],s_j-\underline{(\overline{s_j})}\rangle\\&\overset{\text{(ii)}}{=}(-1)\cdot\|s_j-\underline{(\overline{s_j})}\|^2\\&\overset{\text{(iii)}}{<}0.\end{aligned}$$

[53]We have $\underline{(\overline{s_\theta})}:=(\overline{s_\theta},\overline{s_\theta},\ldots,\overline{s_\theta})^{\mathrm{T}}\in\mathbb{R}^n,\theta=i,j.$

Ad (i). We have $1 = r_{ij} = \frac{\langle s_i - \underline{(\overline{s_i})}, s_j - \underline{(\overline{s_j})}\rangle}{\| s_i - \underline{(\overline{s_i})} \| \cdot \| s_j - \underline{(\overline{s_j})} \|}$.

Ad (ii). It holds $\langle \alpha \cdot x, y \rangle = \alpha \langle x, y \rangle$ for all $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$.

Ad (iii). Since $(I_i, I_j) \in [Q \times Q]_{\mathbf{A}}$, we have $\| s_j - \underline{(\overline{s_j})} \| > 0$.

Case 2 $[(s_j)_{k_2}=1]$. In this case, $0 = \mu - \mu\overline{s_j} + \overline{s_i}$ and $1 = \mu - \mu\overline{s_j} + \overline{s_i}$.

Step 2. There exists an index $k \in \{1,2,\ldots,n\}$ with $(s_i)_k = (s_j)_k = 1$.

Assumption. There is no index $k \in \{1,2,\ldots,n\}$ with $(s_i)_k = (s_j)_k = 1$. Since (I_i, I_j) satisfies **A**, there are indexes $k_1, k_2 \in \{1,2,\ldots,n\}$ with $(s_i)_{k_1} = 0$ and $(s_i)_{k_2} = 1$. Under assumption, $(s_j)_{k_2} = 0$.

Case 1 $[(s_j)_{k_1}=0]$. Then, $0 = \overline{s_i} - \mu\overline{s_j}$ and $1 = \overline{s_i} - \mu\overline{s_j}$.

Case 2 $[(s_j)_{k_1}=1]$. In this case, $0 = \mu - \mu\overline{s_j} + \overline{s_i}$ and $1 = \overline{s_i} - \overline{s_j}$. According to Case 1 of Step 1, this leads to a contradiction.

Step 3. We prove that $\mu = 1$.

According to Step 1 and Step 2, there are indexes $k_1, k_2 \in \{1,2,\ldots,n\}$ with $(s_i)_{k_1} = (s_j)_{k_1} = 0$ and $(s_i)_{k_2} = (s_j)_{k_2} = 1$. Thus, $0 = \overline{s_i} - \mu\overline{s_j}$ and $1 = \mu - \mu\overline{s_j} + \overline{s_i}$. Finally, $1 = \mu + (\overline{s_i} - \mu\overline{s_j}) = \mu + 0 = \mu$.

Step 4. The vectors s_i and s_j are equal.

Since $\mu = 1$, $s_i - \underline{(\overline{s_i})} = s_j - \underline{(\overline{s_j})}$. With Step 1, $\underline{(\overline{s_i})} = \underline{(\overline{s_j})}$. Thus, $s_i = s_j$.

Case 3 $[j < i \wedge (I_j, I_i) \notin \leq]$. Then, $(I_j, I_i) \in <'_{Q|_{\gg}}$. Thus, $0 = b_{ji} = c_{ij}$.

Case 4 $[j < i \wedge (I_j, I_i) \in \leq]$. In this case, (I_j, I_i) belongs to $<'_{Q|_{\cong}}$. Thus,

$$\frac{a_{ji} d_{ji} - b_{ji} c_{ji}}{\sqrt{Z'}} - 1 = 0,$$

with $Z' := [a_{ji} + b_{ji}] \cdot [c_{ji} + d_{ji}] \cdot [a_{ji} + c_{ji}] \cdot [b_{ji} + d_{ji}]$. Since $a_{ji} = a_{ij}$, $b_{ji} = c_{ij}$, $c_{ji} = b_{ij}$, $d_{ji} = d_{ij}$, and $Z' = Z$, we have $\frac{a_{ij} d_{ij} - b_{ij} c_{ij}}{\sqrt{Z}} - 1 = 0$. □

A.7. Proofs in Section 11

Lemma 24. According to Lemma 17, it is the case that

$$\mathrm{CA}(\leq, D^t) = 1 - \frac{2}{m(m-1)} \sum_{k=1}^{4} f_k(\leq, D^t),$$

with[54] $(Z^t := [a^t_{ij} + b^t_{ij}] \cdot [c^t_{ij} + d^t_{ij}] \cdot [a^t_{ij} + c^t_{ij}] \cdot [b^t_{ij} + d^t_{ij}])$

$$f_1(\leq, D^t) := \sum_{(I_i, I_j) \in <'_{Q|^t_{\cong}}} \left[\frac{a^t_{ij} d^t_{ij} - b^t_{ij} c^t_{ij}}{\sqrt{Z^t}} - 1 \right]^2,$$

[54] Entities with (without) superscript t are interpreted under D^t (D). For instance, the entries in a 2×2 table are denoted by a^t_{ij}, b^t_{ij}, c^t_{ij}, d^t_{ij} under D^t, and respectively a_{ij}, b_{ij}, c_{ij}, d_{ij} under D.

$$f_2(\leq, D^t) := \sum_{(I_i,I_j)\in <'_{Q}|^t_{\ll}} \frac{[(n+n_0+n_1)c^t_{ij}]^2}{Z^t},$$
$$f_3(\leq, D^t) := \sum_{(I_i,I_j)\in <'_{Q}|^t_{\gg}} \frac{[(n+n_0+n_1)b^t_{ij}]^2}{Z^t},$$
$$f_4(\leq, D^t) := \sum_{(I_i,I_j)\in <'_{Q}|^t_{\not\asymp}} \frac{[a^t_{ij}d^t_{ij} - b^t_{ij}c^t_{ij}]^2}{Z^t},$$

whereupon

$$<'^t_Q := \{(I_i, I_j)\in Q\times Q : i<j \wedge r_{ij} \text{ exists under } D^t\},$$
$$<'_Q|^t_{\cong} := <'^t_Q \cap \{(I_i, I_j)\in Q\times Q : (I_i, I_j)\in\leq \wedge (I_j, I_i)\in\leq\},$$
$$<'_Q|^t_{\ll} := <'^t_Q \cap \{(I_i, I_j)\in Q\times Q : (I_i, I_j)\in\leq \wedge (I_j, I_i)\notin\leq\},$$
$$<'_Q|^t_{\gg} := <'^t_Q \cap \{(I_i, I_j)\in Q\times Q : (I_i, I_j)\notin\leq \wedge (I_j, I_i)\in\leq\},$$
$$<'_Q|^t_{\not\asymp} := <'^t_Q \cap \{(I_i, I_j)\in Q\times Q : (I_i, I_j)\notin\leq \wedge (I_j, I_i)\notin\leq\}.$$

If r_{ij} exists under D, then it exists under D^t. Thus (+, disjoint set-union):

$$<'_Q|^t_{\cong} = <'_Q|_{\cong} + \delta_{\cong},$$
$$<'_Q|^t_{\ll} = <'_Q|_{\ll} + \delta_{\ll},$$
$$<'_Q|^t_{\gg} = <'_Q|_{\gg} + \delta_{\gg},$$
$$<'_Q|^t_{\not\asymp} = <'_Q|_{\not\asymp} + \delta_{\not\asymp},$$

with (\, set-difference) $\delta_k := <'_Q|^t_k \setminus <'_Q|_k$ for all $k\in\{\cong, \ll, \gg, \not\asymp\}$. Then:

$$f_1(\leq, D^t) = \sum_{(I_i,I_j)\in <'_Q|_{\cong}} \left[\frac{a^t_{ij}d^t_{ij} - b^t_{ij}c^t_{ij}}{\sqrt{Z^t}} - 1\right]^2 + \sum_{(I_i,I_j)\in \delta_{\cong}} \left[\frac{a^t_{ij}d^t_{ij} - b^t_{ij}c^t_{ij}}{\sqrt{Z^t}} - 1\right]^2,$$
$$f_2(\leq, D^t) = \sum_{(I_i,I_j)\in <'_Q|_{\ll}} \frac{[(n+n_0+n_1)c^t_{ij}]^2}{Z^t} + \sum_{(I_i,I_j)\in \delta_{\ll}} \frac{[(n+n_0+n_1)c^t_{ij}]^2}{Z^t},$$
$$f_3(\leq, D^t) = \sum_{(I_i,I_j)\in <'_Q|_{\gg}} \frac{[(n+n_0+n_1)b^t_{ij}]^2}{Z^t} + \sum_{(I_i,I_j)\in \delta_{\gg}} \frac{[(n+n_0+n_1)b^t_{ij}]^2}{Z^t},$$
$$f_4(\leq, D^t) = \sum_{(I_i,I_j)\in <'_Q|_{\not\asymp}} \frac{[a^t_{ij}d^t_{ij} - b^t_{ij}c^t_{ij}]^2}{Z^t} + \sum_{(I_i,I_j)\in \delta_{\not\asymp}} \frac{[a^t_{ij}d^t_{ij} - b^t_{ij}c^t_{ij}]^2}{Z^t}.$$

For $(I_i,I_j)\in Q\times Q$, $a^t_{ij} = a_{ij} + n_1$, $d^t_{ij} = d_{ij} + n_0$, $b^t_{ij} = b_{ij}$, $c^t_{ij} = c_{ij}$. Thus:

$$\begin{aligned} Z^t &:= [a^t_{ij} + b^t_{ij}][c^t_{ij} + d^t_{ij}][a^t_{ij} + c^t_{ij}][b^t_{ij} + d^t_{ij}] \\ &= [(a_{ij} + b_{ij}) + n_1][(c_{ij} + d_{ij}) + n_0][(a_{ij} + c_{ij}) + n_1][(b_{ij} + d_{ij}) + n_0] \\ &\vdots \\ &= Z + R, \end{aligned}$$

with (Note: R=R(n_0,n_1,D), and $R\geq 0$.)

$$\begin{aligned}R = {} & n_0[(a_{ij}+b_{ij})(c_{ij}+d_{ij})(a_{ij}+c_{ij})+(a_{ij}+b_{ij})(a_{ij}+c_{ij})(b_{ij}+d_{ij})]\\ & +n_1[(a_{ij}+b_{ij})(c_{ij}+d_{ij})(b_{ij}+d_{ij})+(c_{ij}+d_{ij})(a_{ij}+c_{ij})(b_{ij}+d_{ij})]\\ & +n_0n_1[(a_{ij}+b_{ij})(c_{ij}+d_{ij})+(a_{ij}+b_{ij})(b_{ij}+d_{ij})+(c_{ij}+d_{ij})(a_{ij}+c_{ij})\\ & +(a_{ij}+c_{ij})(b_{ij}+d_{ij})]+n_0^2n_1[(a_{ij}+b_{ij})(a_{ij}+c_{ij})]+n_0n_1^2[(c_{ij}+d_{ij})(b_{ij}+d_{ij})]\\ & +n_0^2[(a_{ij}+b_{ij})+(a_{ij}+c_{ij})]+n_1^2[(c_{ij}+d_{ij})+(b_{ij}+d_{ij})]+n_0^2n_1^2.\end{aligned}$$

Consider first the term $\sum_{(I_i,I_j)\in <'_Q|_{\cong}}\left[\frac{a^t_{ij}d^t_{ij}-b^t_{ij}c^t_{ij}}{\sqrt{Z^t}}-1\right]^2$. It holds:

$$\begin{aligned}\sum_{(I_i,I_j)\in <'_Q|_{\cong}}\left[\frac{a^t_{ij}d^t_{ij}-b^t_{ij}c^t_{ij}}{\sqrt{Z^t}}-1\right]^2 &= \sum_{(I_i,I_j)\in <'_Q|_{\cong}}\left[\frac{(a_{ij}+n_1)(d_{ij}+n_0)-b_{ij}c_{ij}}{\sqrt{Z+R}}-1\right]^2\\ &\;\vdots\\ &= \sum_{(I_i,I_j)\in <'_Q|_{\cong}}\left[\frac{a_{ij}d_{ij}-b_{ij}c_{ij}}{\sqrt{Z}}-1\right]^2+F_1,\end{aligned}$$

whereupon F_1=F_1(n_0,n_1,D,$\leq$) is defined as

$$F_1 = \sum_{(I_i,I_j)\in <'_Q|_{\cong}}\left[G_1\left[2\left(\frac{a_{ij}d_{ij}-b_{ij}c_{ij}}{\sqrt{Z}}-1\right)+G_1\right]\right],$$

with G_1=G_1(n_0,n_1,D),

$$G_1 = [a_{ij}d_{ij}-b_{ij}c_{ij}]\left[\frac{1}{\sqrt{Z+R}}-\frac{1}{\sqrt{Z}}\right]+\frac{n_0n_1+n_0a_{ij}+n_1d_{ij}}{\sqrt{Z+R}}.$$

For the remaining terms, in analogy to the previous term (left to the reader):

$$\begin{aligned}\sum_{(I_i,I_j)\in <'_Q|_{\ll}}\frac{[(n+n_0+n_1)c^t_{ij}]^2}{Z^t} &= \sum_{(I_i,I_j)\in <'_Q|_{\ll}}\frac{[nc_{ij}]^2}{Z}+F_2,\\ \sum_{(I_i,I_j)\in <'_Q|_{\gg}}\frac{[(n+n_0+n_1)b^t_{ij}]^2}{Z^t} &= \sum_{(I_i,I_j)\in <'_Q|_{\gg}}\frac{[nb_{ij}]^2}{Z}+F_3,\\ \sum_{(I_i,I_j)\in <'_Q|_{\neq}}\frac{[a^t_{ij}d^t_{ij}-b^t_{ij}c^t_{ij}]^2}{Z^t} &= \sum_{(I_i,I_j)\in <'_Q|_{\neq}}\frac{[a_{ij}d_{ij}-b_{ij}c_{ij}]^2}{Z}+F_4,\end{aligned}$$

with appropriate F_i=$F_i(n_0,n_1,D,\leq)$ for $2\leq i\leq 4$. With these terms,

$$f_1(\leq, D^t) = f_1(\leq, D) + \left[F_1 + \sum_{(I_i,I_j)\in\delta_{\cong}} \left[\frac{a^t_{ij}d^t_{ij} - b^t_{ij}c^t_{ij}}{\sqrt{Z^t}} - 1\right]^2\right],$$

$$f_2(\leq, D^t) = f_2(\leq, D) + \left[F_2 + \sum_{(I_i,I_j)\in\delta_{\ll}} \frac{[(n+n_0+n_1)c^t_{ij}]^2}{Z^t}\right],$$

$$f_3(\leq, D^t) = f_3(\leq, D) + \left[F_3 + \sum_{(I_i,I_j)\in\delta_{\gg}} \frac{[(n+n_0+n_1)b^t_{ij}]^2}{Z^t}\right],$$

$$f_4(\leq, D^t) = f_4(\leq, D) + \left[F_4 + \sum_{(I_i,I_j)\in\delta_{\not\asymp}} \frac{[a^t_{ij}d^t_{ij} - b^t_{ij}c^t_{ij}]^2}{Z^t}\right].$$

Defining (Note: F_i^*=$F_i^*(n_0,n_1,D,\leq)$ for all $1\leq i\leq 4$.)

$$F_1^* := F_1 + \sum_{(I_i,I_j)\in\delta_{\cong}} \left[\frac{a^t_{ij}d^t_{ij} - b^t_{ij}c^t_{ij}}{\sqrt{Z^t}} - 1\right]^2,$$

$$F_2^* := F_2 + \sum_{(I_i,I_j)\in\delta_{\ll}} \frac{[(n+n_0+n_1)c^t_{ij}]^2}{Z^t},$$

$$F_3^* := F_3 + \sum_{(I_i,I_j)\in\delta_{\gg}} \frac{[(n+n_0+n_1)b^t_{ij}]^2}{Z^t},$$

$$F_4^* := F_4 + \sum_{(I_i,I_j)\in\delta_{\not\asymp}} \frac{[a^t_{ij}d^t_{ij} - b^t_{ij}c^t_{ij}]^2}{Z^t},$$

we obtain functional relationships for the functions f_i, $1\leq i\leq 4$:

$$f_i(\leq, D^t) = f_i(\leq, D) + F_i^*(n_0, n_1, D, \leq).$$

Finally, we have

$$\begin{aligned}
\mathrm{CA}(\leq, D^t) &= 1 - \frac{2}{m(m-1)} \sum_{k=1}^{4} f_k(\leq, D^t) \\
&= \left[1 - \frac{2}{m(m-1)} \sum_{k=1}^{4} f_k(\leq, D)\right] - \frac{2}{m(m-1)} \sum_{k=1}^{4} F_k^*(n_0, n_1, D, \leq) \\
&= \mathrm{CA}(\leq, D) - F^*(n_0, n_1, D, \leq),
\end{aligned}$$

with $F^*(n_0, n_1, D, \leq) := \frac{2}{m(m-1)} \sum_{k=1}^{4} F_k^*(n_0, n_1, D, \leq)$. □

References

Airasian, P.W., Bart, W.M., 1973. Ordering theory: a new and useful measurement model. Educational Technology 13, 56–60.

Bart, W.M., Krus, D.J., 1973. An ordering-theoretic method to determine hierarchies among items. Educational and Psychological Measurement 33, 291–300.

Birkhoff, G., 1937. Rings of sets. Duke Mathematical Journal 3, 443–454.

Bishop, Y.M.M., Fienberg, S.E., Holland, P.W., 1975. Discrete Multivariate Analysis: Theory and Practice. MIT Press, Cambridge, MA.

Bortz, J., 1989. Statistik. Springer, Berlin.

Cramér, H., 1946. Mathematical Methods of Statistics. Princeton Univ. Press, Princeton, NJ.

Dayton, C.M, Macready, G.B., 1976. A probabilistic model for validation of behavioral hierarchies. Psychometrika 41 (2), 189–204.

Doignon, J.-P., Falmagne, J.-Cl., 1985. Spaces for the assessment of knowledge. International Journal of Man–Machine Studies 23, 175–196.

Doignon, J.-P., Falmagne, J.-Cl., 1999. Knowledge Spaces. Springer, Berlin.

Fischer, G., 1995. Lineare Algebra. Vieweg, Braunschweig/Wiesbaden.

Goodman, L.A., Kruskal, W.H., 1954. Measures of association for cross classifications. Journal of the American Statistical Association 49, 732–764.

Goodman, L.A., Kruskal, W.H., 1959. Measures of association for cross classifications: II. Further discussion and references. Journal of the American Statistical Association 54, 123–163.

Goodman, L.A., Kruskal, W.H., 1963. Measures of association for cross classifications: III. Approximate sampling theory. Journal of the American Statistical Association 58, 310–364.

Goodman, L.A., Kruskal, W.H., 1972. Measures of association for cross classifications: IV. Simplification of asymptotic variances. Journal of the American Statistical Association 67, 415–421.

Guttman, L., 1941. An outline of the statistical theory of prediction. In: Horst, P., et al. (Eds.), (Ed.), The Prediction of Personal Adjustment. Bulletin, vol. 48. Social Science Research Council, New York, pp. 253–318.

Held, T., Korossy, K., 1998. Data analysis as a heuristic for establishing theoretically founded item structures. Zeitschrift für Psychologie 206, 169–188.

Held, T., Schrepp, M., Fries, S., 1995. Methoden zur Bestimmung von Wissensstrukturen—Eine Vergleichstudie. Zeitschrift für Experimentelle Psychologie 42 (2), 205–236.

Kendall, M.G., Stuart, A., 1973. The Advanced Theory of Statistics, vol. II. Hafner, New York.

Leeuwe, J.F.J.V., 1974. Item Tree Analysis. Nederlands Tijdschrift voor de Psychologie 29, 475–484.

Liebetrau, A.M., 1983. Measures of Association. Sage Publications, Beverly Hills, CA, USA.

Pearson, K., 1904. On the theory of contingency and its relation to association and normal correlation. Drapers' Co. Memoirs. Biometric Series, vol. 1. London.

Schrepp, M., 1999. On the empirical construction of implications between bi-valued test items. Mathematical Social Sciences 38, 361–375.

Schrepp, M., 2003. A method for the analysis of hierarchical dependencies between items of a questionnaire. Methods of Psychological Research-Online 19, 43–79.

Schrepp, M., Held, T., Albert, D., 1999. Component-based construction of surmise relations for chess problems. In: Albert, D., Lukas, J. (Eds.), (Eds.), Knowledge Spaces: Theories, Empirical Research, Applications. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 41–66.

Tschuprow, A.A., 1918/1919/1921. On the mathematical expectation of the moments of frequency distributions. Biometrika 12/13, 140–169, 185–210, 283.

Wesiak, G., Albert, D., Ünlü, A., 2004. How to generate and validate hypotheses on surmise relations between tests. Exemplified for inductive reasoning tests. Manuscript, University of Graz, Graz, Austria/Europe.

Yule, G.U., Kendall, M.G., 1950. An Introduction to the Theory of Statistics. Charles Griffin, London.