# Editorial: Promising Reasoning Test Ideas not yet published

*Klaus D. Kubinger*

University of Vienna

***Preamble****:*
Over a quarter of a century ago, when I focused my research work on a renaissance of *Raymond B. Cattell*'s so-called *Objective Personality tests* (cf. Kubinger, 1997, 2006, 2009), there was not only a (German-edited) test-battery on the market (though hardly in practical use; Häcker, Schmidt, Schwenkmezger, & Utz, 1975), which offered about 50 test concepts (indicating by numeration that at least 197 such concepts were developed in the original English version); but there was also the famous compendium by Cattell and Warburton (1967), which summarized some empirical evidence of those test concepts. Although most of them were not examined according to relevant quality criteria of psychological test construction, and most notably, the test-scores were never standardized based on a representative sample, there was a seemingly endless quantum of promising test concepts at an interested researcher's disposal. A similar situation arises for interested researchers in a research area from which I am withdrawing, now: the dimensionality of *reasoning* – though not with respect to the incomparable number of *Cattell*'s test concepts. It seems worthwhile to issue some conceptualized reasoning tests that have partly demonstrated their psychometric quality, particularly with respect to their empirical adequacy of measurement. That is, there are several tests with publication potential, which, however, need further in-depth research investment. This special issue may serve as a stimulation for this purpose.

**Author Note**

Prof. Klaus D. Kubinger, PhD, MSc. c/o University of Vienna, Faculty of Psychology. klaus.kubinger@univie.ac.at

## Introduction

Since *Louis L. Thurstone* (cf. Thurstone, 1938), the factor *reasoning* is well-established within intelligence theories. Accordingly, psychological assessment defines reasoning as the "ability to realize regularities and logically compelling connections in order to put them to appropriate use" (Kubinger, 2019, p. 244; translation by the author). There are of course many tests in practical application, largely based on *John C. Raven*'s test concept of so-called *(Progressive) Matrices*.

There are, however, hardly any psychological matrices tests at a practitioners' disposal that meet psychometric standards, which even *Raven's Standard Progressive Matrices* (SPM) has proven not to do (cf. Kubinger, Formann, & Farkas, 1991). Nevertheless, there is the *Viennese Matrices-Test* (WMT-2; Formann, Waldherr, & Piswanger, 2011), which does not only fulfill the psychometric requirement of fitting the Rasch model – take into account that scoring (just) the number of solved items, as is very common with psychological tests, urgently calls for Rasch model's validness (cf. Fischer, 1995); but the *Viennese Matrices-Test* is (probably as the first one) also constructed by using item generating rules: That is, if the regularities of an item's components are discovered by the testee, his/her appropriate further use of them automatically leads to the solution. As a consequence, content validity of the test is immediately given, as testees who solve many items have proven to master the underlying logical rules to a high extent. Furthermore, in applying for the Rasch model fitting items *Gerhard H. Fischer*'s LLTM (linear logistic test model; Fischer, 1973; see also Fischer, 2005, as well as Kubinger, 2008), Formann (1973) established for each item generating rule the respective difficulty parameter, so that by adequate combination of the rules new items can be composed of whichever item difficulty the examiner chooses (see Fischer & Pendl, 1980, for the technical implementation of this idea, in particular within adaptive testing). It is regrettable that the *Viennese Matrices-Test* is only published for the German speaking market, though obviously non-verbal administration is possible.

Apart from matrices tests, there are other test conceptualizations that refer to the above given definition of reasoning. We systematize in the following two times three categories of reasoning tests.

With reference to *Raymond B. Cattell* (Cattell, 1963), it can be differentiated between *fluid* and *crystallized* intelligence – and therefore between *fluid* and *crystallized* reasoning. Traditionally, due to *Cattell*'s efforts to obtain culture-free tests, the *fluid* facet is given more focus than the *crystallized* one in the construction of reasoning tests. With reference to *Adolf O. Jäger* (Jäger, 1984), it can be differentiated between *verbal* (better: *lexical*), *numerical*, and *figural* contents, as concerns intelligence test items in general. Crossing both *fluid* vs. *crystallized* facets and *lexical* vs. *numerical* vs. *figural* contents leads to six categories of reasoning tests. Of course, all of them seem worthwhile for the elaboration of separate test conceptualizations. That is because in practice, when consulting a specific case, it is often not requested to assess a client's reasoning ability in general, but rather with reference to a specific (occupational) profile

of requirements. For instance, applying a test measuring *crystallized* reasoning by means of *lexical* contents often seems more appropriate than applying a test measuring *fluid* reasoning by means of *figural* contents (i.e. a matrices test, as an example).

A large state facility for public safety initiated the corresponding development of tests. Accordingly, several conceptualizations were created.

## Contents

Appropriate conceptualizations are summarized in Figure 1, according to the indicated six categories of reasoning tests.

|  | crystallized | fluid |
|---|---|---|
| lexical | *Family-Relation Reasoning-Test* | *Reality-contradicting Syllogisms* |
| numerical | *Equations* | *Numerical Topologies* |
| figural | *Culture-referenced Pictographic Analogies* | *Two-way Figural Reasoning-Test* |

**Figure 1:**

*Six reasoning test conceptualizations introduced in this special issue: two different intelligence facets and three different contents*

In this special issue, a *crystallized-lexical* reasoning test is suggested (Poinstingl & Sparfeldt, 2023), which has been presented numerous times at several conferences, and some master theses have already proven its potential usefulness. The task is to find the type of family relation between two short story's protagonists, deduced from other family members' relations, which are explicitly described in the story. This is an original conceptualization, never before suggested by any other authors.

In contrast, a test concept using the formal-logical means of a "syllogism" – two premises unequivocally imply a certain conclusion – has been tried before, particularly by Srp (1994): Apart from applying adaptive tailored testing (see e.g. Kubinger, 2016), the test *Syllogisms* was the first (and up to now, only) test using the so-called sequential response format; that is, the answer options are presented one after the other as long as the testee decides the given option is wrong, instead of the simultaneous presentation of all answer options at once as is commonly used with multiple choice items. By doing this, the probability of lucky guessing is substantially reduced. As this test was withdrawn from the market due to the closure of the concerned publishing house, some efforts for a renewal are presented in this special issue (Treiber & Kubinger, 2023). In particular, the items were redesigned according to several logically completely irrelevant components (i.e. the conclusion is always in reality not true and the premises as well as the conclusion use sometimes the conjunctive mood), and include one more distractor, created by reformulating one of the three logically

wrong conclusions. Obviously, that test concept refers to the *fluid* facet, as education-based knowledge is hardly of relevance though *lexical* contents are used.

At least with regard to the answer format, the test conceptualization of using mathematical equations with two unknowns is original (Kubinger & Gamsjäger, 2023). The task is to find both values of the unknowns that fulfill the given equation. For this, a special multiple choice answer format is used: „2 × (1 out of 4)", that is, four options are offered for each unknown. Of course, the required ability refers to the *crystallized* facet concerning *numerical* contents.

The *fluid-numerical* test conceptualization referring to objects' position and arrangement in the space, which has to be logical to continue, is also original (Kubinger & Heuberger, 2023). Although the objects are digits (i.e. mainly single digits, seldom two- or more digit numbers) the *crystallized* facet is hardly needed for the solution.

Following the well-known tests asking for verbal (i.e. lexical) analogies, a non-lexical application has been conceptualized using pictures of everyday objects (Kubinger, Ünal, & Schnait, 2023). That is, reasoning is examined in a *figural* sense, where the looked-for relation between the depicted objects is based on *crystallized* power.

Furthermore, a type of a matrices test is dealt with by Bartok and Kubinger (2023) in order to examine the *fluid* facet with *figural* contents. What makes it new and unique is, that a) the (5×5-) matrices do not present all elements but one, however instead present only a selection of them which are non-redundant for finding the element in question (i.e. the solution); b) the element in question is neither always located on the same position nor always on the position "last row/last column".

When the conceptualization of the discussed six reasoning tests is presented in the following contributions of this special issue, not only are typical items illustrated but the applied item generating rules are also demonstrated. Empirical analyses are always given, how in far the constructed item-pool meets psychometric standards, that is whether the Rasch model holds – and in case some items have to be deleted in order to fit the model, suggestions are made on how the concerned item generating rules should be better revised. However, the main concern from the point of view of intelligence theory is in which way these six reasoning tests correlate with each other and hence, eventually find common factors. This is dealt with by Kubinger (2023) in this special issue.

## References

Bartok, L. & Kubinger, K. D. (2023). Conceptualization of a new "Two-way Figural Reasoning-Test". *Psychological Test and Assessment Modeling, 65*, 321-338.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*, 1-22.

Cattell, R. B., & Warburton, F. W. (1967). *Objective personality and motivation tests*. Champaign (IL): University of Illinois Press.

Fischer, G. H. (1973). The linear logistic test model as an instrument in educational research. *Acta Psychologica, 37*, 359-374.

Fischer, G. H. (1995). Derivations of the Rasch Model. In G. H. Fischer & I. W. Molenaar (eds.), *Rasch models* (pp. 15–38). New York: Springer.

Fischer, G. H. (2005). Linear logistic test models. *Encyclopedia of Social Measurement, 2*, 505-514.

Formann, A. (1973). *Die Konstruktion eines neuen Matrizentests und die Untersuchung des Lösungs-verhaltens mit Hilfe des linearen logistischen Testmodells* [The construction of a new matrices-test and the analysis of the testees' solving behaviour by the linear logistic test-model]. Unveröffentliche Dissertation [unpublished doctoral thesis], Universität Wien [University of Vienna].

Formann, A. K., Waldherr, K., & Piswanger, K. (2011). *Wiener Matrizen-Test 2, WMT-2* [Viennese Matrices-test, WMT-2]. Göttingen: Hogrefe.

Häcker, H., Schmidt, L. R., Schwenkmezger, P., & Utz, H. E. (1975). *Objektive Testbatterie, OA-TB 75* [Objective test-battery, OA-TB 75]. Weinheim: Beltz.

Jäger, A. O. (1984). Intelligenzstrukturforschung: Konkurrierende Modelle, neue Entwicklungen, Perspektiven [Research on the intelligence structure: competing models, new progresses and outlooks]. *Psychologische Rundschau*, *35*, 21-35.

Kubinger, K. D. (1997). Zur Renaissance der objektiven Persönlichkeitstests sensu R.B. Cattell [The renaissance of R. B. Cattell's objective personality tests]. In H. Mandl (ed.), *Bericht über den 40. Kongreß der Deutschen Gesellschaft für Psychologie in München 1996* [Report of the 40th congress of the German society of psychology, in Munich 1996] (pp. 755-761). Göttingen: Hogrefe.

Kubinger, K. D. (2006). Ein Update der Definition von Objektiven Persönlichkeitstests: Experimentalpsychologische Verhaltensdiagnostik [An updated definition of objective personality tests: experiment-based assessment of behavior]. In T. M. Ortner, R. Proyer & K. D. Kubinger (eds.). *Theorie und Praxis Objektiver Persönlichkeitstests* [Theory and Practice of objective personality tests] (pp. 38-52). Bern: Huber.

Kubinger, K. D. (2008). On the revival of the Rasch model-based LLTM: From constructing tests using item generating rules to measuring item administration effects. *Psychology Science Quarterly, 50,* 311-327.

Kubinger, K. D. (2009). The technique of objective personality-tests sensu R. B. Cattell nowadays: The Viennese pool of computerized tests aimed at experiment-based assessment of behavior. *Acta Psychologica Sinica, 41*, 1024-1036.

Kubinger, K. D. (2016). Adaptive testing. In K. Schweizer & C. DiStefano (eds.), *Principles and methods of test construction* (pp. 104-119). Göttingen: Hogrefe.

Kubinger, K. D. (2019). *Psychologische Diagnostik – Theorie und Praxis psychologischen Diagnostizierens* (3rs ed.) [*Psychological Assessment – Theory and Practice of Psychological Consulting*]. Göttingen: Hogrefe.

Kubinger, K. D. (2023). On the dimensionality of reasoning. *Psychological Test and Assessment Modeling, 65*, 437-447.

Kubinger, K. D., Formann, A.K. & Farkas, M.G. (1991). Psychometric shortcomings of Raven´s Standard Progressive Matrices (SPM) in particular for computerized testing. *European Review of Applied Psychology, 41*, 295-300.

Kubinger, K. D. & Gamsjäger, C. (2023). Conceptualization of the reasoning-test "(Mathematical) Equations". *Psychological Test and Assessment Modeling, 65*, 339-353.

Kubinger, K. D. & Heuberger, N. (2023). Conceptualization of the reasoning-test "Numerical Topologies". *Psychological Test and Assessment Modeling, 65*, 354-387.

Kubinger, K. D., Ünal, S., & Schnait, B. (2023). Conceptualization of the reasoning-test "Culture-referenced Pictographic Analogies". *Psychological Test and Assessment Modeling, 65*, 388-401.

Poinstingl, H. & Sparfeldt, J. R. (2023). Conceptualization of the "Family Relations Reasoning Test" (FRRT). *Psychological Test and Assessment Modeling, 65*, 414-436.

Srp, G. (1994). *Syllogismen* [Syllogisms]. Frankfurt/M.: Swets Test Services.

Thurstone, L. L. (1938). *Primary mental abilities.* Chicago: University of Chicago Press.

Treiber, S. & Kubinger, K. D. (2023). Conceptualization of the reasoning-test "Reality-contradicting Syllogisms". *Psychological Test and Assessment Modeling, 65*, 402-413.