Learning and Individual Differences 109 (2024) 102401

Contents lists available at ScienceDirect

Learning and Individual Differences

journal homepage: www.elsevier.com/locate/lindif

Unscrambling creativity measurement: An invitation to better formalize the domain generality-specificity of creativity with psychometric modeling

Nils Myszkowski [a,*], Martin Storme [b], Pinar Çelik [b]

[a] Department of Psychology, Pace University, New York, NY, USA
[b] IESEG School of Management, Univ. Lille, CNRS, UMR 9221 - LEM - Lille Economie Management, Lille, France

ARTICLE INFO

Keywords:
Creativity
Individual differences
Measurement
Psychometrics

ABSTRACT

A key question in creativity research concerns whether creativity is best understood as domain-specific or domain-general (Baer & Kaufman, 2005). Yet, the literature is plagued with insufficiently defined notions of domain specificity and generality of creativity. For example, researchers have alternatively discussed domain-generality as either a positive manifold (i.e., positive interdomain correlations), or as a unidimensional factor model (e.g., Ivcevic, 2007), without acknowledging the difference or providing a justification for their particular choice. We argue that there are diverse ways to conceptualize domain generality and specificity, with different implications (Kan et al., 2019). We propose to unscramble confusions in the literature by defining taxonomy of models that imply a general factor, domain specific factors, both, or neither. We review statistical frameworks used to test and compare them, examine their implications, and invite creativity researchers to clarify how they conceptualize and formalize generality and specificity in their research.

Statement of educational relevance: A clear representation of how creativity is structured across domains is fundamental to the understanding of how creativity is to be learned and trained. Notably, the existence of a general creativity implies that creativity might be learned and trained in general (i.e., across multiple domains), while domain specificity implies that it might be learned primarily at the domain (or subdomain) level. Beyond the traditional framing of this question as a dichotomy or continuum of generality-specificity, we argue that there are in fact many types of conceptual and statistical models for understanding the structure of creativity across domains, which have different implications.

1. Introduction

One of the great debates in creativity research concerns the structure of creativity across domains: Is creativity domain general or domain specific? For example, does a person's musical creativity tell us about their literary or scientific creativity? Is there such a personal characteristic as general creativity? Our aim with the current paper is not so much to directly address this question – there are already comprehensive and up to date accounts of the state of theory and empirical research on the topic (Baer, 2015; Kaufman et al., 2017). Instead, we want to discuss lacks and discrepancies in the definition of the problem itself. More specifically, we argue that domain generality – and thus necessarily its (seemingly) opposite proposition, domain specificity – are not sufficiently defined, and that this lack of clarity prevents well-defined hypotheses to be tested empirically and discussed theoretically. Our objective is threefold: 1) to pinpoint the lack of clarity in definitions of domain generality and specificity, 2) to propose a clearer approach by defining by a set of empirically testable models that represent different theoretical hypotheses, and 3) to discuss frameworks that may allow to test these models and compare them. In the remainder of this introduction, we will present first the domain generality-specificity debate and its relevance to creativity research.

1.1. *One of the great debates of creativity research*

The domain-generality-specificity debate is a key debate in creativity research. We could probably trace its origin in the first research that considered whether creative and artistic abilities are separate from other forms of mental ability (e.g., Binet, 1908). Debates on the domain generality of creativity have been linked from the start to similar debates on the general factor of intelligence. For example, the (hypothetical) general factor of intelligence is commonly referred to as *g*, and the

* Corresponding author.
E-mail address: nmyszkowski@pace.edu (N. Myszkowski).

https://doi.org/10.1016/j.lindif.2023.102401
Received 30 March 2023; Received in revised form 25 October 2023; Accepted 9 December 2023

(hypothetical) general factor of creativity is sometimes referred to as *c* (Baer, 2012). In intelligence research, debates on the existence of *g* have also been central for over a century (e.g., Spearman, 1904) and are still vivid (Van Der Maas et al., 2006).

More generally, the question of whether a general psychological attribute can explain most of the behaviors of a domain (e.g., academic achievement, creativity, job performance) is central to many fields of psychology. For example, if we examine the debate on general intelligence, we find that the main theoretical implication of a potential general factor is its very study and discussion. Indeed, if there is no general intelligence, there is no study to be made of it, and the term itself might disappear, to the benefit of a study of other specific attributes – those referred to as primary (mental) abilities in the famous Thurstone-Spearman debate (Spearman, 1939; Thurstone, 1938). In other words, debates on the existence of a general factor shape how a domain of individual attributes is studied (how we study its relations with its predictors, how we study its outcomes, how tests are constructed and analyzed, etc.).

1.2. Implications for measurement

Much like in other fields, the domain general-specific debate has important implications for the measurement of creativity. On a practical level first, the domain generality of creativity is a base assumption in many general models of personality or vocational interests. For example, measures of openness often comprise facets or communalities of items referring to individual characteristics such as imaginativeness and aesthetic appreciation (Christensen et al., 2019), but these facets, although more specific than openness, remain general in terms of creativity domains (e.g., there is no musical imagination or literature appreciation facet). Likewise, Holland's (1997) model of vocational interests discusses and measures one's Artistic (A) inclination in general. If there is no such thing as general creativity (or general artistic creativity), then such measures would be measuring an inexistant construct. Second, some degree of domain generality is assumed in most measures of creativity themselves (e.g., Barron & Welsh, 1952; Mednick, 1968; Torrance, 1966), which do not necessarily measure creativity in general, but still often purport to measure at least some relatively broad domain of creativity (e.g., figural creativity, musical creativity).

Other researchers have regarded general creativity as a collection of behaviors that are viewed – and perhaps only socially constructed (Westmeyer, 1998) – as creative. This is the idea that creativity may exist as a domain of behavior, but not as an explanatory variable. This view has important implications for measurement, as both classical and modern test theory relies on the causal theory of measurement, which would then be rejected. General creativity measurement (or perhaps scoring is here a more appropriate term) might still however be possible, through using some form of weighted composite score (e.g., Christensen & Golino, 2021).

However, from a conceptual viewpoint, the most common psychometric procedures (e.g., factor analysis, item response theory, sum/average scoring, Cronbach's α) are based on the causal theory of measurement – i.e., reflective measurement – according to which latent, unobserved person attributes are assumed to cause behaviors (i.e., item scores) (Markus & Borsboom, 2013). Consequently, if such an attribute as general creativity is hypothesized to exist as an explanation for creative behaviors, then it is conceptually consistent to attempt to build measurement devices that will allow to infer a person's general creativity. Such instruments would require generating items with minimal domain specificities or that cover domains broadly enough – either through some form of latent variable modeling (e.g., traditional factor analysis, item-response theory modeling), or through the classical test theory set of axioms (i.e., through sum/average scoring). Conversely, if we do not assume such an explanatory factor to exist, attempting to estimate a person's general creativity becomes conceptually inconsistent; we would be attempting to measure an attribute that we posit does not exist. Thus, one might abandon measuring general creativity and measure creativity at the level(s) at which one posits that a common factor exists instead.

We shall note here that a number of psychometric and conceptual advances – such as the bifactor model (Holzinger & Swineford, 1937; Reise, 2012), multi-trait multi-method models, and latent state-trait models (Steyer et al., 1999) – allow general/stable aspects of an attribute to be measured while specific/state aspects are accounted for. These methods have notably been used in creativity research (Barbot et al., 2016; Zandi et al., 2022), which has demonstrated that creativity in general can be modeled and estimated, while domain effects are also modeled and estimated (and controlled for). From these different approaches, we can conclude that, if the presence of a general creativity latent explanatory variable may be a necessary assumption to measure general creativity, it does not imply that domain specificities cannot be accounted for and measured within the same model.

1.3. Implications for trainings and intervention research

Beyond measurement, there are also implications of the general-specific debate for research and practice. As Baer (2012) noted, creativity trainings are largely based on the assumption that general creativity exists. Indeed, creativity trainings usually focus on a limited number of domains, and therefore usually assume that the enhanced skills resulting from the training will permeate into creativity in other aspects. Yet, if creativity tasks or domains are independent, then attempts to train creativity as a whole would instead require to provide a training for every possible domain of creativity. Further, the inquiry of how individual differences, motivational aspects and environmental factors influence creativity would be impacted. If a general creativity factor exists, then these factors could be studied as having effects on creativity in general, which then would impact creative behaviors. But, if general creativity does not exist, then the individual and situational factors that play a role in creativity would in fact need to be studied separately for each domain. The specialization-differentiation hypothesis (Barbot & Tinio, 2015) has for example been presented as an alternative explanation for correlations between some creativity tasks, that does not rely on a general creativity factor, but instead, on the activation of skills from specializing in a domain. According to this hypothesis, resources specific to a particular creative domain can, to some extent, transfer to domains with similar task demands, which would explain positive correlations across creativity domains without necessitating the existence of a general creativity factor.

1.4. A multilevel and complex question

The generality-specificity debate is relevant at different levels of creativity. For example, zooming in on domains and subdomains, one may discuss if there is such an attribute as general creativity, as general artistic creativity, as general musical creativity, as general jazz music creativity, or as general piano jazz music creativity. Certainly, the generality-specificity debate receives most of its spotlight when it comes to a general *c* factor, but it is in fact just as relevant (although perhaps not as important) at different levels of inquiry within the field of creativity psychology. The Amusement Park Theory (APT; Baer & Kaufman, 2005), for example, provides a description for how these different levels (general creativity requirements, general thematic areas, domains, micro-domains and tasks) are articulated and interweaved.

We shall note here that researchers (Plucker & Beghetto, 2004) have rejected the dichotomy and advocated for a more hybrid position between generality and specificity, on the account that, in any case, the components involved in creativity themselves have domain general and domain specific aspects. Consequently, the problem of generality-specificity would then not appear as a dichotomy, but rather as a continuum from fully domain general (i.e., no specific factors) to fully domain specific (i.e., no general factor). Further, it has been advanced

(Stevenson et al., 2021) that creativity fundamentally requires intelligence (which can be used in several domains) and expertise (which is domain specific), which explains how creativity is probably both general and specific at the same time.

The problem however is that even if we accept that creativity may best be trained or measured as both general and specific (and thus decide that the general/specific debate is an archaic dichotomy that should be avoided), this does not automatically lead to optimal training strategies, nor measurement methods. For this, some kind of theory-based decision on how creative potential and behaviors are conceptually structured is crucial. That is, an explicit measurement model is vital (Borsboom, 2006). We argue that, a non-decision, albeit conceptually reasonable, remains hardly operationalizable for measurement and research purposes.

2. The confusing definitions of domain generality-specificity

As mentioned above, to optimally study the causes and antecedents of creativity and/or its specific domains requires a theory-based decision on whether the items/tasks measured in a study should be conceptualized and statistically modeled as domain general or domain specific. It goes without saying that different decisions on conceptualization and modeling can result in different conclusions. Unfortunately, as we will see, researchers are often not very explicit in such decisions and they often implicitly use various underlying assumptions for how domain generality and specificity in creativity manifests itself in the data structure. In what follows, we will discuss how in the literature creativity researchers hold various implicit definitions of domain generality and specificity, and how those definitions guide their research approach.

In most contemporary research in creativity that discusses the issue, domain generality is defined through its prediction of "positive correlations among the levels of creativity exhibited by individuals in different domains" (Baer, 2012, p. 19). Domain specificity being considered as the opposite of domain generality, it then predicts the opposite, meaning no (or low) correlations (Baer, 2012; Ivcevic, 2007). Evidently, this definition of domain generality is based on the eventual empirical consequences of what is defined. Indeed, if there is such a thing as domain general creativity, then it would imply that creative behaviors should be correlated across domains, and therefore observing creative behaviors being correlated across domains would indicate the existence of domain creativity. For this reason, in the remainder of the article, we refer to this as a *consequentialist* definition of domain-generality/specificity. The consequentialist approach defines domain generality as a *positive manifold* (Bartholomew et al., 2009; Van Der Maas et al., 2006) phenomenon. The positive manifold, a term generally used in research on *g*, refers to the presence of (positive) correlations among tasks/items/domains. Here, translating to creativity, the positive manifold would manifest itself as correlations between creativity measurements across domains. Domain general creativity is – according to the consequentialist definition discussed in this section – defined as that which produces (or would produce) a *positive manifold* of creative behaviors.

One explanatory hypothesis for a positive manifold phenomenon is the existence of a common explanatory factor of the tasks/items/domains (although we will later explain that it is only one of the possible explanations). However, a positive manifold may also appear from one or several common causes that we important during the development of a particular expertise but cease to be important over time. Consider how general cognitive abilities are essential for acquiring domain-specific knowledge and skills. However, once an individual achieves a certain level of expertise, the importance of these general abilities in fostering creativity could diminish. Then, a positive manifold of creative behaviors may appear because of how the expertise was acquired, but these behaviors may not have a substantial common cause anymore.

Nevertheless, we note that most researchers are trained – if not hard-wired – whenever they see a positive manifold (generally represented in a correlation matrix with mostly positive correlations between items/tasks), to proceed to some application of a common factor model, which assumes that an unobserved (i.e. latent) attribute causes responses to the tasks or items. This assumption is relatively explicit in the factor analytic and the item response theory modeling traditions, but is in fact also made, albeit implicitly, when using more routine procedures, such as Cronbach's alpha or sum/average scoring. In fact, this assumption is so omnipresent that it is frequently assumed (again, not necessarily explicitly) without prior testing (Borsboom, 2006). Like many psychometric decisions (Borsboom, 2006; Myszkowski & Storme, 2019), this "default" assumption is generally made to satisfy practical imperatives (e.g., computing scores, estimating reliability) – and this is certainly not specific to creativity research.

Although it is hard to trace its origins, it is possible that the consequentialist definition of domain generality is more of an involuntary approximation than an intentional definition. Summarizing the question of domain generality to positive correlations might simply be the symptom of a more general reluctance of researchers to understand and investigate the structure of tests and test batteries, perhaps due to some discomfort in certain psychometric techniques – which is not specific to creativity research (Borsboom, 2006) – or some concern regarding their outcomes (e.g., poor fit indices) or feasibility (e.g., low sample size, software unavailability). This is unfortunate, because, as we will next discuss, a positive manifold often suggests that there is *some structure* underlying the task scores, but it does not (at least directly) show precisely what the exact nature of this structure is. It may be the common factor model, but not necessarily.

More recently, a number of researchers have been more explicitly using a factor analytic approach to the problem, to directly identify the statistical cause of potential correlations between creativity tasks (e.g., Qian et al., 2019). Since this approach specifies domain generality as a common explanatory attribute that causes creative behaviors, this could be called a *causalist* definition of domain general creativity. In this approach, if domain-general creativity exists, one should be able to identify a single explanatory attribute that explains a complete set of creative behaviors (i.e., item scores). In other words, evidence for domain generality is tied here more explicitly to whether a general latent attribute realistically causes creative behaviors – which may be approached theoretically, empirically, or both. It is effectively a direct application of the causal theory of measurement. In practice, evidence to support domain generality would be found through some method of factor analysis (e.g., exploratory factor analysis, confirmatory factor analysis, item-response theory modeling) indicating that a single factor is a plausible explanation for creative behaviors across domains.

Unfortunately, although the factor analytic approach in theory allows testing richer conceptualizations of creativity, a lot of the research that addresses domain generality / specificity of creativity tends to test for the presence of a single latent factor structure – possibly comparing such a structure with a correlated domain factors model – disregarding potentially more complex structures, such as structures that combine domain specificity with domain generality (i.e., bifactor models). Although more in line with our proposition to use a modeling approach, studies often fail to consider the full spectrum of possibilities. For example, second order factor models (e.g., Kapoor et al., 2021) and/or bifactor models (e.g., Kapoor et al., 2021; McKay et al., 2017) may not be investigated. The use of a narrow scope of models in the exploration of the structure of creativity may restrict the identification of potentially important factors or dimensions that contribute to creative behaviors, and lead to skewed conclusions and psychometric applications.

2.1. The problem of equating a positive manifold with domain generality

In the literature, there is an almost automatic equating of the positive manifold with a common factor, with both definitions often used interchangeably. For example, Ivcevic's definition (2007, p. 272), which is commonly referred to (Baer, 2012), bundles together the positive manifold ("high intercorrelations") and the common factor(s) hypothesis ("common set of psychological descriptors"). Certainly, the existence of a common explanatory attribute (i.e., the causal definition) implies a positive manifold (i.e., the consequentialist definition): as Baer points out (2012, p. 19), "the theory that creativity is domain-general therefore predicts positive correlations among the levels of creativity exhibited by individuals in different domains". In other words, a *c* domain-general factor would, indeed, lead to a positive manifold.

So, where is the problem in interchanging the two definitions? In a nutshell, the existence of a positive manifold does not necessarily imply a general factor (Van Der Maas et al., 2006). In fact, a positive manifold may be observed both in the presence as well as in the absence of a general factor. Therefore, if a domain general factor is inferred directly from a positive manifold, then several alternate (and often plausible) explanations for the positive manifold are ignored. For example, several common latent factors may exist (instead of one) that may themselves be correlated (i.e., a correlated-factors model), leading to a positive manifold – a researcher observing positive correlations might in this case be looking at the manifestation of several specific factors.

Another possible issue is that researchers may find themselves in the seemingly incongruent situation of observing a positive manifold but failing to observe a general factor. In this situation, a researcher or reader might either conclude that creativity is domain specific (because of the absence of a common factor), or conclude that creativity is domain general (because of the positive manifold), thus creating two contradictory interpretations of one result. Further, a positive manifold implies that creative tasks/domains influence one another, while a common factor model implies that the general factor is not caused by its manifestations. Therefore, if one is to design a training, a positive manifold would imply that training a domain would tend to impact the others, while, assuming a common factor model, one would only alter creativity in other domains if they were training the general factor itself (and, consequently, its domains). Thus, the two definitions make different implications as to how trainings shall be designed or might work when they do.

In addition, confusion regarding what defines domain generality (and its counterpart, specificity) leads to confusion about what an empirical result indicates or not. The stricter causal definition of domain generality invites researchers to conclude to research specificity more often than the more permissive consequentialist definition, because the common factor model is one among several explanations of a positive manifold. Finally, both the causalist and consequentialist definition insufficiently specify how general creativity (if any), domain specific factors and tasks are configured, and thus what they represent. For example, in the causalist definition (where general creativity is represented as common factor), if one assumes that creativity is both general and specific, then one could imagine that the general factor is a common cause for domain-specific creativities, which themselves cause creative behaviors/items (e.g., McKay et al., 2017). But, alternatively, one could imagine that domain-factors cause behaviors/items, while general creativity (e.g., Myszkowski & Storme, 2021) – or perhaps specific abilities that themselves are caused by a common factor of creativity (e.g., Barbot et al., 2016) – directly cause(s) behaviors/items. In this example, the scores derived from these different approaches are not only different, but they also have different interpretations (DeMars, 2013). In the consequentialist definition, similar uncertainties occur. For example, if domain generality is retained per this definition, then one would expect a positive manifold, but should it be observed at the behavior/item level, or at the domain level?

3. Expanding the general-specific dichotomy: towards a modeling approach

A possible solution would be for researchers to settle on a clear approach for what is specifically meant by domain generality, and what constitutes its evidence. In line with recent research involving various modeling approaches to generality and specificity (e.g., Barbot et al., 2016; McKay et al., 2017), we argue that there are different ways to articulate domain generality and specificity, beyond a dichotomy or a continuum between generality and specificity. Instead, we argue that we should clearly define a set of plausible structural models that are empirically testable and comparable, and that represent relations (and non-relations) between observed tasks (i.e., measured creative behaviors) and various factors (specific and general). Hopefully, evidence in favor of one or some of these models may help us more clearly articulate the relations between domain general creativity (if any), domain-specific factors and observable behaviors.

Before presenting the model taxonomy, we shall note that, while we consider that we present here a clearer set of models that are plausible and interpretable than has been presented before, using statistical modeling alone to make decisions on the structure of a set of tasks is itself imperfect, notably for the reason that different models can be equivalent (for example, in structural equation modeling, a model with two correlated specific factors would be equivalent statistically to a hierarchical model). Thus, we want to clarify that we do not argue that statistical modeling itself is the only way to explore the structure of creativity (or of particular domains/subdomains), even though we think that it provides a clearer way to describe and investigate the structure of creativity than general-specific dichotomy.

Below, we present these different models conceptually. The presentation of the different models assumes a usual measurement situation for creativity, with a number of tasks (i.e., item scores, behaviors) possibly indicated by domain-general creativity and/or specific factors. Importantly, although we aim to provide an adaptable set here, but this set does not represent all measurement paradigms. For example, there may be situations involving rater effects, or more nested levels of generality-specificity to consider (e.g., general creativity, domain, subdomain and item), which we do not discuss here.

To note, we present here *structural* models, not response models. Depending on the item characteristics, these models may notably be empirically examined from a traditional (i.e. linear) factor analytic framework (in general, using a structural equation modeling package), an item-response theory framework (in general, using an item response theory package) and/or general purpose latent variable modeling software. Also, we do not discuss the distribution of the latent variables involved. For simplification, and without loss of generality, we will assume tasks that yield continuous item scores that are well approximated by normal distributions, that potentially indicate latent factors that are themselves normally distributed, and we assume that all relations between variables are linear. To note, for several instruments used in creativity research, this approximation may be largely inaccurate – such as, for example, for binary response scales (e.g., Mednick, 1968), for count responses (e.g., Myszkowski & Storme, 2021), or for some ordinal scales (e.g., Carson et al., 2005) – which may impact the feasibility of estimating a number of these models.

There is nothing to be inferred from the order of presentation of the models, although we attempted to describe them from the simplest to the most complex. The models are presented (in relatively minimal forms, showing few items and domains) as path diagrams in Fig. 1. In all diagrams, task scores (e.g., an originality score at a figural divergent thinking task) are noted X_1, X_2, etc., general creativity (if any) is noted C, and domain facets or specific factors (if any) are noted D', D'', etc.

3.1. Models with no specific or general creativity factor

We will first discuss some models that do not assume the existence of latent explanatory variables for task scores.

3.1.1. Independent tasks model (ITM)

One first plausible model is a model where the item scores are completely independent from one another, and are not explained by any

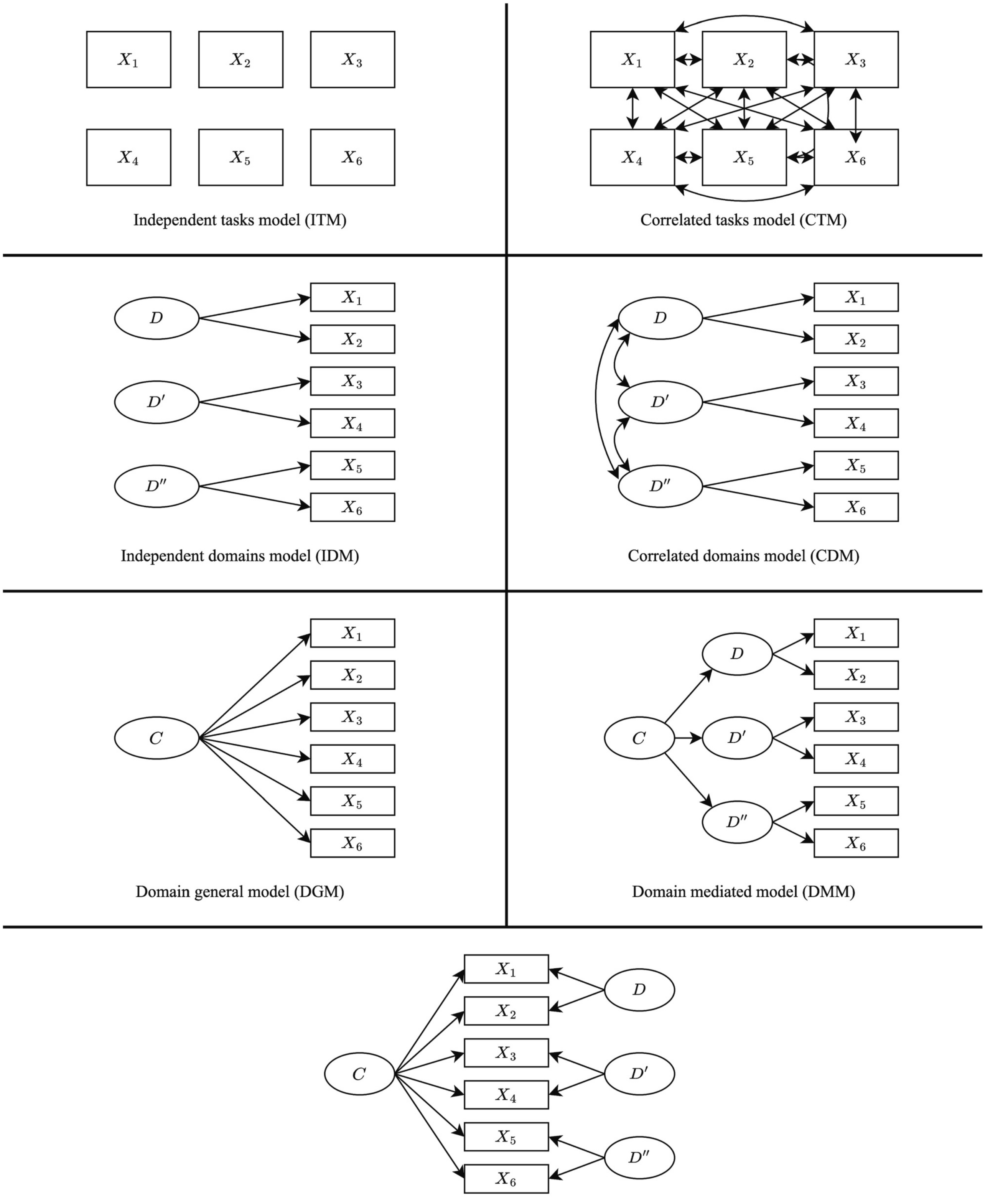

Fig. 1. Diagrams representing the classification of models.

common factor – specific or general. This model implies no effect from a general factor or specific factor, and thus does not assume the existence of any explanatory latent variable. This model is often a model of interest in confirmatory factor analysis, for the reason that it is traditionally used as baseline model (i.e., worst model possible) for incremental fit indices (e.g., comparative fit index, Tucker-Lewis index).

3.1.2. *Correlated tasks model (CTM)*

In this model, all item responses are directly related to one another, and no item is explained by a specific or general factor. Again here, this model does not assume the existence of any explanatory latent variable. However, in comparison with the previous model, this model suggests that a person's creativity for a given tasks affects (directly) their creativity for all other tasks. This is also a model of interest in confirmatory factor analysis, in which it is the saturated model (i.e., best model possible, perfect fit), and the distance to this model is used for the computation of absolute fit indices (e.g., χ^2 test, root mean square error of approximation).

3.1.3. Variations

We may imagine variations of this model. Notably, we may imagine a model in which certain task scores are related to one another, while some are not. In this model, a person's creativity for a given task may affect their creativity for some other tasks. Depending on the type of item, various methods may be used to estimate empirically which relations between tasks may be considered as spurious and which may not, and various methods may be used to interpret the outcome of this relation filtering process. In general, this set of tools falls under the category of network psychometric modeling (see Isvoranu et al., 2022 for an introduction).

3.2. Models with only specific factors

In this section, we present models that assume the existence of specific domain effects, represented by explanatory latent variables, but no domain general factor.

3.2.1. Independent domains model (IDM)

In this model, task scores are explained by a set of latent explanatory variables that represent specific factors, while the scores are locally independent (i.e., uncorrelated). The specific factors are themselves uncorrelated as well. Under the consequentialist definition, this model would be representative of domain specificity.

3.2.2. Correlated domains model (CDM)

Like previously, in this model, task scores are explained by a set of latent variables that represent specific factors with locally independent scores. Unlike previously, however, the specific factors are assumed to correlate with one another, instead of being independent. Under the consequentialist definition, this model, if it accurately represents some data, would present evidence in favor of domain generality. It would however not be sufficient if using a causalist definition of domain generality.

3.2.3. Variations

These models may be extended, notably by including certain local dependencies between tasks, which would typically be identified either from theory, using modification indices in structural equation modeling, or using a network model approach to specify a network of residuals (Epskamp et al., 2017). In addition, certain domains may be correlated while some are not, and similarly, these relations may be identified from theory (e.g., we hypothesize that certain domains are related because they are similar in some way), using modification indices, or using a network model approach (Epskamp et al., 2017). Finally, some domains may be thought to potentially affect tasks that are theoretically from another domain (e.g., a verbal domain factor could affect a scientific creativity task that involves writing a scientific report), and these relations (generally referred to as cross-loadings) may also be specified.

3.3. Models with a general factor

Finally, we present here a set of models that imply a general factor, whether they imply the existence of domain specific factors or not.

3.3.1. Domain-general model (DGM)

In this model, task scores are all, across domains, directly explained by a single explanatory latent variable. Under the causal definition of domain generality, this model, if empirically supported, supports domain-generality.

3.3.2. Domain-mediated model (DMM)

This model is a hierarchical model, in which specific domains explain the tasks, and are themselves explained by a general (second-order) factor. This model, if empirically supported, suggests the existence of a general factor, which predicts task scores indirectly through domain factors. Thus, here, domains serve as *mediators* for the effect of the general factor. In personality research terms, we could refer to domains here as facets of creativity.

3.3.3. Domain-controlled model (DCM)

This model is essentially a bifactor model (Holzinger & Swineford, 1937; Reise, 2012) (also discussed as a multidimensional compensatory model in item response theory) where all tasks are explained by a common general factor, while each task is also explained by a specific domain factor. This model, if empirically supported, indicates the existence of both a general factor and specific factors (under the causalist definition). Here, the general factor directly causes the task scores, while domain-specific factors are statistically controlled for (as opposed to being mediators in the previously discussed model). In other words, domains are represented as *extraneous* (i.e., nuisance, shared variance) variables.

3.3.4. Variations

Here, several variations may be imagined for these models. First, like explained before, one may explore potential local dependencies between the tasks themselves. Second, some correlations between specific factors may be explored as well. Third, cross-loadings may be used to represent domains that might influence the task of another domain. Like discussed previously, these modifications may be justified theoretically or be the result of empirical explorations (modification indices, exploratory factor analyses, network models, etc.).

4. Comparing models empirically: frameworks and software

Although we do not aim to provide a tutorial on estimating these models and comparing them empirically, we will briefly discuss how this might be achieved. First, it appears relatively clear, from the path diagrams and the description, that a number of the models presented here include latent variables. Thus, the estimation of all of these models requires a modeling framework that can estimate latent variables. In addition, several latent variables need to be specified in some of these models. Further, because it is probably unrealistic – particularly in creativity measurement (Myszkowski & Storme, 2021) – to assume that latent variables will predict with equal strength all of their indicators (items or domains), a modeling framework that allows for freely estimated factor loadings (in classical test theory this is often referred to as a congeneric model, while in item-response theory the terms generalized and non-Rasch models are more common) is also required. At this point, candidate frameworks essentially include (generalized) structural equation modeling and item-response theory – which can be seen as different approaches to the same class of models (e.g., Mellenbergh, 1994).

In general, item-response theory software tend to be more focused on (large scale) test construction, and therefore it is frequent that, in this approach, a unidimensional model is aimed at and/or assumed from the start, or that the focus is more the choice of an appropriate response function (e.g., one-parameter logistic vs. two-parameter logistic) than the number of factors and their structure (which is the focus here). Thus, unfortunately, many item response theory modeling software packages – with exceptions (e.g., Chalmers, 2012) – are limited in the structures that can be investigated, and they also often tend to limit the possibility to add local dependencies. Conversely, the structural equation modeling framework tends to focus more on factor structure and often easily allows to add local dependencies. Unfortunately, this may come at the cost of not always having many options in terms of item response distributions and models, beyond linear/gaussian and logistic/binomial (Myszkowski, 2021).

In spite of item response limitations, the (generalized) structural equation modeling framework easily allows to estimate and compare models, using various estimation methods and fit indices or comparison statistics, depending on the item distribution – in many cases (Gaussian and categorical outcomes), the "full range" of fit indices (comparative fit index, root mean squared error of approximation, etc.) is available, although fewer indices may be available for other distributions (e.g., Myszkowski & Storme, 2021). Once all models are specified, they can therefore be compared and ranked, and conclusions can be drawn from these model comparisons.

We shall note that some of the variations of the models presented here require to explore and decide on retaining certain local dependencies (or correlations between factors) while discarding others (i.e., creating a sparse network of local dependencies). While there are some tools in structural equation modeling (and item response theory) that may be used here, such as using modification indices to respecify a model stepwise, it risks overfitting datasets. Currently, this process of selecting and discarding relations between variables is much more developed in the network psychometrics framework (Isvoranu et al., 2022). Fortunately, the frameworks of latent variable modeling and network modeling tend to be more and more compatible (see Kan et al., 2020, for example).

5. Discussion

Over the years, research on the domain generality of creativity has yielded mixed results. While a number of studies indicate that domain general cognitive processes, such as general mental ability (Kim, 2008), and domain general traits, such as openness (Feist, 1998), tend to positively predict creativity across domains, other studies have found that creative behaviors tend to not be well summarized with a single factor (e.g., Kaufman, 2012). Certainly, acquiring sufficiently wide-ranging evidence that would accurately quantify to what extent creativity is general and to what extent it is specific is not easily feasible across all possible domains of creativity. However, our approach has important applications at the more local levels that many creativity researchers are interested in (e.g., at the level of a domain, a subdomain, or a particular study). What we offer is that in such applications researchers can consider a more complete array of possible models that conceptualize and formalize statistically the relations between different creative outcomes. Even though the scope of one particular study at a more local level of creativity may not allow much generalization, such studies can still advance the field by using a clear measurement model, and over time, with cumulative evidence, could still inform theory formation about how creativity is structured. We hope to have demonstrated that the array of possible models goes beyond discussing the general-specific dichotomy or continuum and has concrete applications for creativity researchers.

We advance here that the definition of what constitutes a case of domain generality or a case of domain specificity is unclear. The main reason, in our opinion, is that it is unclear whether a positive manifold (i.e., positively correlated creative behaviors) is a sufficient condition for domain generality, or whether evidence for an underlying (causal) latent variable (i.e., a general *c* creativity factor) is necessary. Importantly, the former may exist without the latter (Kan et al., 2019; Van Der Maas et al., 2006), while creativity researchers regularly use the two arguments interchangeably. Further, as we are discussing simultaneously domain-general creativity (i.e., *c*), domain-specific factors and creative behaviors themselves, the combination of these levels implied by the domain-general and the domain-specific hypothesis is vague. For example, does the domain-general hypothesis indicate that domain-specific factors are themselves caused by *c*, or does it indicate that *c* causes creative behaviors directly, controlling for domain specificities?

We discussed how to solve these ambiguities by expanding the traditional general-specific dichotomy into a series of clearly defined, testable and comparable models, which represent different ways that one may interpret domain generality and/or domain specificity. We first defined models that do not assume the existence of domain general or specific factors – creative behaviors, directly reinforcing one another (independent tasks model, or ITM) or not (dependent tasks model, or DTM). Then, we discussed models that assume the existence of domain specific factors but not of a domain general factor – the independent domains model (IDM) and the correlated domains model (CDM). Finally, we discussed models with a domain general factor (*C*), where either the general factor explains creative behaviors with no domain specific factor (domain-general model, or DGM), the general factor explains creative behaviors through the domains (domain-mediated model, or DMM), or the general factor explains the creative behaviors, controlling for domain-specific factors (domain-controlled model, DCM). We then discuss how one may approach the comparison of these models, and notably explain the various requirements for testing these hypotheses. Currently, (generalized) structural equation modeling is likely the most flexible approach to test these different models, although network psychometrics might provide additional tools to explore variations of these models (especially regarding local dependencies and correlations between latent variables), while multidimensional item response theory might provide more options in terms of item response models.

Although researchers have discussed why creativity is probably partly both general and specific, they have mainly discussed how different levels of creativity are related to with one another conceptually (Baer & Kaufman, 2005; Plucker & Beghetto, 2004), but have not proposed how to translate these propositions into testable models. In other words, admitting creativity as both domain general and specific does not provide us with a clear measurement model that we can use to estimate (and thus, predict, and train) creativity and its domains. Although we did not solve this question here, we hope to have presented a clearer array of possibilities. While many of these models may be seen as compromises between domain generality and specificity, we think that the psychometric approach to the question which we provide here does not really use a sliding scale of domain generality vs. specificity. Instead, we present the two (generality and specificity) as not even mutually exclusive propositions, as our classification indicates that there are some models that support domain generality (in our opinion, these would simply be the models where there is a general factor), some models that support domain specificity (these would be the models with domain-specific factors), some models that support both simultaneously (with different roles of specific domains, as controls or mediators), and some models that do not support either (these would be the models with no general nor specific factors). For example, the psychometric model ultimately selected by Kaufman (2012) indicates 5 domain specific factors explaining item scores and no general factor, which corresponds to the independent domains model.

While we presented different types of models that can be used to study the structure of creativity across domains, it can be argued that, since both causalist and consequentialist definitions have been used and interchanged for domain generality (Ivcevic, 2007), and since these different definitions lead to generality being supported by a different set of models (for example, a correlated domains model presents domain generality per the consequentialist definition, but not per the causalist definition), then discussions (on a given dataset or in general) regarding the structure of creativity across domains should perhaps avoid the terminology of general and specific. Instead, we suggest that discussions and empirical investigations are more precisely specified as measurement models. It can be further argued that the terminology used in our model taxonomy, because it still employs "general" and "specific", remains flawed.

Although we present an approach to the general-specific question that uses various structural models, it should be noted that this approach has some caveats that researchers should be mindful about. First, in many situations, some models may be easier to estimate on a given dataset than others. Notably, bifactor models (i.e., a domain-controlled model here) frequently fail to converge in estimation, or lead to anomalous results that can cause interpretation issues (Eid et al., 2017). Second, there are several situations in which two types of models that are theoretically different could nevertheless be statistically equivalent and produce the same fit indices. This would be the case, for example, for a domain-general model with three tasks versus a correlated tasks model with the same three tasks. These two models would produce the same fit indices. Thus, there are situations in which two different theoretical models may not be empirically comparable. Further, it is probable that, in a number of datasets, several of the proposed models provide a good fit (e.g., if nested, they may not significantly differ in fit, and if non nested, they may have similar fit indices). In such situations, from a purely statistical – that is, data driven, perspective – one would have to retain several models, with possibly contradictory interpretations. Further, some models presented, because of their higher complexity, will necessarily have better fit than others (for example, the correlated domains model cannot be outperformed by the domain general model in terms of chi-square in a structural model). This calls for an approach to model fit that penalizes model complexity in some way, and that would therefore not automatically favor more complex models. Finally, beyond the models differing in structure, the interpretation of the latent factors that are estimated differs as a function of the model considered. Notably, specific factors in a bifactor model have a different interpretation than facets in a second-order (i.e., domain general model). Therefore, the meaning of what a general factor or what a domain factor is depends on the model considered.

Another limitation of the approach presented here is that it focuses on creativity as generalizable (or not) across domains of application, but it does not address how different abilities and skills may be configured. For example, we did not discuss how abilities like divergent thinking and intelligence, as well as personality traits and acquired skills from expertise may be structured, nor if the same structure of abilities, traits and skills would be expected to manifest across domains. Thus, in a sense, we only present how a person's creative *products* may be structured, leaving aside how a person's creative *resources* may be structured. In a given model, it is however possible to assume a particular configuration for creative skills, traits and/or abilities in a particular way (e.g., with a hierarchical structure), while domain effects are controlled for (e.g., Barbot et al., 2016).

As we pointed, general-specific debates may occur at different levels in creativity research. The taxonomy of models presented here may therefore be used differently in different contexts. For example, in a dataset with only artistic creativity tasks, we may consider artistic creativity to be a general factor, and subdomains of artistic creativity (music, painting, etc.) to be domains. Alternatively, in a dataset with artistic and non-artistic creativity tasks, we may consider artistic creativity to be a domain. In situations where the two approaches are possible, researchers should carefully define and clarify what is intended to be the level of the general factor and the level of the domain factors on the basis of theory.

Whether our classification is sufficiently exhaustive or not, we think that research on the topic still needs to progress in its methods and its theory. Regarding its methods, we propose that future research shall test a more extensive pool of candidate structural models, beyond simply acknowledging the presence or absence of correlations between items, or the goodness of fit of a unidimensional model. Regarding theory, much like in intelligence research (Kan et al., 2019), it appears to us that a mutualist perspective, where creative behaviors (or domains) would (at least to some extent) directly reinforce one another and create a positive creativity manifold – as opposed to being caused by a general c factor – is a perspective that remains largely unexplored, and that may explain why some creative domains or some creative behaviors might form clusters, without necessarily being caused by some more or less general latent attributes.

Declaration of competing interest

The authors declare no conflict of interest.

References

Baer, J. (2012). Domain specificity and the limits of creativity theory. *The Journal of Creative Behavior, 46*(1), 16–29. https://doi.org/10.1002/jocb.002
Baer, J. (2015). *Domain specificity of creativity*. Academic Press.
Baer, J., & Kaufman, J. C. (2005). Bridging generality and specificity: The amusement park theoretical (APT) model of creativity. *Roeper Review, 27*(3), 158–163. https://doi.org/10.1080/02783190509554310
Barbot, B., Besançon, M., & Lubart, T. (2016). The generality-specificity of creativity: Exploring the structure of creative potential with EPoC. *Learning and Individual Differences, 52*, 178–187. https://doi.org/10.1016/j.lindif.2016.06.005
Barbot, B., & Tinio, P. P. L. (2015). Where is the "g" in creativity? A specialization–differentiation hypothesis. *Frontiers in Human Neuroscience, 8*. https://www.frontiersin.org/articles/10.3389/fnhum.2014.01041.
Barron, F., & Welsh, G. S. (1952). Artistic perception as a possible factor in personality style: Its measurement by a figure preference test. *The Journal of Psychology, 33*(2), 199–203. https://doi.org/10.1080/00223980.1952.9712830
Bartholomew, D. J., Deary, I. J., & Lawn, M. (2009). A new lease of life for Thomson's bonds model of intelligence. *Psychological Review, 116*(3), 567–579. https://doi.org/10.1037/a0016262
Binet, A. (1908). La psychologie artistique de Tade Styka. *L'Année Psychologique, 15*(1), 316–356. https://doi.org/10.3406/psy.1908.3760
Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika, 71*(3), 425–440. https://doi.org/10.1007/s11336-006-1447-6
Carson, S. H., Peterson, J. B., & Higgins, D. M. (2005). Reliability, validity, and factor structure of the creative achievement questionnaire. *Creativity Research Journal, 17*(1), 37–50. https://doi.org/10.1207/s15326934crj1701_4
Chalmers, R. P. (2012). Mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software, 48*(1), 1–29. https://doi.org/10.18637/jss.v048.i06
Christensen, A. P., Cotter, K. N., & Silvia, P. J. (2019). Reopening openness to experience: A network analysis of four openness to experience inventories. *Journal of Personality Assessment, 101*(6), 574–588. https://doi.org/10.1080/00223891.2018.1467428
Christensen, A. P., & Golino, H. F. (2021). On the equivalency of factor and network loadings. *Behavior Research Methods, 53*(4), 1563–1580. https://doi.org/10.3758/s13428-020-01500-6
DeMars, C. E. (2013). A tutorial on interpreting Bifactor model scores. *International Journal of Testing, 13*(4), 354–378. https://doi.org/10.1080/15305058.2013.799067
Eid, M., Geiser, C., Koch, T., & Heene, M. (2017). Anomalous results in G-factor models: Explanations and alternatives. *Psychological Methods, 22*(3), 541–562. https://doi.org/10.1037/met0000083
Epskamp, S., Rhemtulla, M., & Borsboom, D. (2017). Generalized network psychometrics: Combining network and latent variable models. *Psychometrika, 82*(4), 904–927. https://doi.org/10.1007/s11336-017-9557-x
Feist, G. J. (1998). A Meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review, 2*(4), 290–309. https://doi.org/10.1207/s15327957pspr0204_5
Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments*. Psychological Assessment Resources.
Holzinger, K. J., & Swineford, F. (1937). The bi-factor method. *Psychometrika, 2*(1), 41–54. https://doi.org/10.1007/BF02287965
Isvoranu, A.-M., Epskamp, S., Waldorp, L., & Borsboom, D. (2022). *Network psychometrics with R: A guide for behavioral and social scientists*. Routledge.
Ivcevic, Z. (2007). Artistic and everyday creativity: An act-frequency approach. *The Journal of Creative Behavior, 41*(4), 271–290. https://doi.org/10.1002/j.2162-6057.2007.tb01074.x
Kan, K.-J., de Jonge, H., van der Maas, H. L. J., Levine, S. Z., & Epskamp, S. (2020). How to compare psychometric factor and network models. *Journal of Intelligence, 8*(4), 35. https://doi.org/10.3390/jintelligence8040035
Kan, K.-J., van der Maas, H. L. J., & Levine, S. Z. (2019). Extending psychometric network analysis: Empirical evidence against *g* in favor of mutualism? *Intelligence, 73*, 52–62. https://doi.org/10.1016/j.intell.2018.12.004
Kapoor, H., Reiter-Palmon, R., & Kaufman, J. (2021). Norming the muses: Establishing the psychometric properties of the Kaufman domains of creativity scale. *Journal of Psychoeducational Assessment, 39*. https://doi.org/10.1177/07342829211008334, 073428292110083.
Kaufman, J. C. (2012). Counting the muses: Development of the Kaufman Domains of Creativity Scale (K-DOCS). *Psychology of Aesthetics, Creativity, and the Arts, 6*(4), 298–308. https://doi.org/10.1037/a0029751
Kaufman, J. C., Glăveanu, V. P., & Baer, J. (2017). *The Cambridge handbook of creativity across domains*. Cambridge University Press.
Kim, K. H. (2008). Meta-analyses of the relationship of creative achievement to both IQ and divergent thinking test scores. *The Journal of Creative Behavior, 42*(2), 106–130. https://doi.org/10.1002/j.2162-6057.2008.tb01290.x

Markus, K. A., & Borsboom, D. (2013). Reflective measurement models, behavior domains, and common causes. *New Ideas in Psychology, 31*(1), 54–64. https://doi.org/10.1016/j.newideapsych.2011.02.008

McKay, A. S., Karwowski, M., & Kaufman, J. C. (2017). Measuring the muses: Validating the Kaufman Domains of Creativity Scale (K-DOCS). *Psychology of Aesthetics, Creativity, and the Arts, 11*(2), 216–230. https://doi.org/10.1037/aca0000074

Mednick, S. A. (1968). The remote associates test. *The Journal of Creative Behavior, 2*(3), 213–214. https://doi.org/10.1002/j.2162-6057.1968.tb00104.x

Mellenbergh, G. J. (1994). Generalized linear item response theory. *Psychological Bulletin, 115*(2), 300–307. https://doi.org/10.1037/0033-2909.115.2.300

Myszkowski, N. (2021). Development of the R library "jrt": Automated item response theory procedures for judgment data and their application with the consensual assessment technique. *Psychology of Aesthetics, Creativity, and the Arts, 15*(3), 426–438. https://doi.org/10.1037/aca0000287

Myszkowski, N., & Storme, M. (2019). Judge response theory? A call to upgrade our psychometrical account of creativity judgments. *Psychology of Aesthetics, Creativity, and the Arts, 13*(2), 167–175. https://doi.org/10.1037/aca0000225

Myszkowski, N., & Storme, M. (2021). Accounting for variable task discrimination in divergent thinking fluency measurement: An example of the benefits of a 2-parameter Poisson counts model and its Bifactor extension over the Rasch Poisson counts model. *The Journal of Creative Behavior, 55*(3), 800–818. https://doi.org/10.1002/jocb.490

Plucker, J. A., & Beghetto, R. A. (2004). Why creativity is domain general, why it looks domain specific, and why the distinction does not matter. In *Creativity: From potential to realization* (pp. 153–167). American Psychological Association. https://doi.org/10.1037/10692-009.

Qian, M., Plucker, J. A., & Yang, X. (2019). Is creativity domain specific or domain general? Evidence from multilevel explanatory item response theory models. *Thinking Skills and Creativity, 33*, Article 100571. https://doi.org/10.1016/j.tsc.2019.100571

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47*(5), 667–696. https://doi.org/10.1080/00273171.2012.715555

Spearman, C. (1904). "General intelligence," objectively determined and pmeasured. *The American Journal of Psychology, 15*(2), 201–292. https://doi.org/10.2307/1412107

Spearman, C. (1939). Thurstone's work re-worked. *Journal of Educational Psychology, 30*(1), 1.

Stevenson, C., Baas, M., & van der Maas, H. (2021). A minimal theory of creative ability. *Journal of Intelligence, 9*(1), 9. https://doi.org/10.3390/jintelligence9010009

Steyer, R., Schmitt, M., & Eid, M. (1999). Latent state–trait theory and research in personality and individual differences. *European Journal of Personality, 13*(5), 389–408. https://doi.org/10.1002/(SICI)1099-0984(199909/10)13:5<389::AID-PER361>3.0.CO;2-A

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

Torrance, E. P. (1966). *The Torrance tests of creative thinking–norms—Technical manual research edition—Verbal tests, forms a and B—Figural tests*. Forms A and B: Personnel Press.

Van Der Maas, H. L. J., Dolan, C. V., Grasman, R. P. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review, 113*(4), 842–861. https://doi.org/10.1037/0033-295X.113.4.842

Westmeyer, H. (1998). The social construction and psychological assessment of creativity. *High Ability Studies, 9*(1), 11–21. https://doi.org/10.1080/1359813980090102

Zandi, N., Karwowski, M., Forthmann, B., & Holling, H. (2022). How stable is the creative self-concept? A latent state-trait analysis. *Psychology of Aesthetics, Creativity, and the Arts*. https://doi.org/10.1037/aca0000521. No Pagination Specified-No Pagination Specified.