A Confirmatory Factor Analysis Approach to Test Anxiety

Peter F. Halpin*, Cibele da-Silva+ & Paul De Boeck*

* University of Amsterdam  + University of Brasilia

This research was supported by a postdoctoral grant from the National Science and Engineering Council of Canada.

PFH is now at New York University; PDB is now at Ohio State University

Correspondence Information:
Dr. Peter F. Halpin
New York University
246 Greene St., 3 floor
New York NY
10013-6677
e-mail: peter.halpin@nyu.edu
Abstract

This paper addresses the role of test anxiety in aptitude testing. Our approach is rooted in confirmatory factor analysis (CFA). We find that the usual parameter constraints used for model identification in CFA have non-trivial implications for the effects of interest. We suggest two methods for dealing with this identification problem. Firstly, we consider testable parameter constraints that identify the proposed model. Secondly, we consider structural relations that do not depend on model identification. In particular we derive the partial factor correlation between a test and an external variable, conditional on test anxiety, and show that this correlation (a) is not affected by the choice of model identification constraints and (b) can be estimated using true score theory.

Keywords: test anxiety, confirmatory factor analysis, model identification, partial correlation.
A Confirmatory Factor Analysis Approach to Test Anxiety

Test anxiety is anxiety that originates from the appraisal of being tested as threatening (Zeidner, 1998). In the modern day context of high-stakes testing, it has become increasingly important to understand the relationship between test anxiety and test performance (e.g., Cizek & Burg, 2006; van der Embse & Hasson, 2012; Weems et al., 2010). In particular, it is important to (a) assess whether and to what extent test scores systematically underestimate the ability of anxious test takers, and (b) find a means of “correcting” test scores for the effects of test anxiety. In this paper we address these two issues by providing a measurement model of ability that incorporates test anxiety.

A related line of research concerns the criterion validity of intelligence tests (Reeve & Bonaccio, 2008; Reeve, Heggestad, & Lievens, 2009; Wicherts & Zand Scholten, 2010). It has been shown that test anxiety can bias the observed correlation between intelligence tests and their criteria, although the direction of the bias is a matter of ongoing debate. In general, this research highlights the importance of test anxiety as a confounding variable in correlation studies involving aptitude tests. We address this issue by modelling the structural relations among ability, test anxiety, and external variables.

Our approach is rooted in confirmatory factor analysis (CFA). We find that test anxiety presents unconventional challenges related to model identification and therefore leads to some relatively general methodological considerations. In particular, we argue that tests can not be treated as unidimensional in the presence of test anxiety, but must be regarded as (possibly) two-dimensional. Unlike similar multidimensional models (e.g., the bi-factor model, multi-trait multi-method models), it can not be assumed that ability and test anxiety are orthogonal. Consequently we are faced with a specific type of rotational indeterminacy, which we refer to as confound indeterminacy.

We propose two main ways of dealing with confound indeterminacy. Firstly, we consider testable parameter constraints that serve to eliminate it. Depending on the
research context, these constraints may be viewed either as simplifying assumptions placed on the model or as substantive hypotheses. Once the model is identified, the contributions of ability and test anxiety to test performance can be disentangled, leading to a measure of ability that is not biased by test anxiety.

Secondly, we consider structural relationships that remain invariant under confound indeterminacy. Specifically, we show that the partial factor correlation between ability and an external variable, conditional on test anxiety, is invariant and can be estimated using the methods of true score theory. This result is useful for researchers who are committed to using observed correlations, but does not allow for interpretation of other structural relationships in the model.

In contrast to recent approaches (e.g., da-Silva & Gomes, 2011; Wicherts & Zand Scholten, 2010), we focus on the role of test anxiety as a continuous latent variable rather than its role as a manifest grouping variable. The use of a grouping variable leads naturally to considerations about measurement invariance and differential item functioning. The use of a continuous factor allows us to address these same issues as well as structural relations among tests, test anxiety, and other variables. Moreover, our approach is congenial to the growing literature on the assessment of test anxiety via self-report questionnaires (e.g., Cizek & Burg, 2006; Oostdam & Meijer, 2003; Wren & Benson, 2004). In keeping with the usual practice of CFA, we phrase the discussion in terms of continuous indicators (e.g., test scores) that are linearly related to their measurement constructs. Extension to discrete indicators with non-linear measurement equations (e.g., item level analysis) can be made by means of the familiar arguments (e.g., Takane & de Leeuw, 1987; McCullagh & Nelder, 1989). We do not consider non-linear structural relations (e.g., latent moderation).

It should be emphasized that this paper presents a methodological solution to a substantive problem. It is therefore neither exclusively methodological nor exclusively applied in its scope. Because the problem motivates novel research in CFA, it will be of interest to methodologists. However, we have not attempted to present our results in anything like their highest generality, which is a task for further research. Our presentation is specific to the context of test anxiety, and we illustrate the application
our results using an example from educational testing.

The remainder of this paper is organized as follows. The next section reviews the literature on test anxiety in order to develop an appropriate CFA model. The subsequent section demonstrates that the proposed model is subject to confound indeterminacy and considers parameter constraints that eliminate this indeterminacy. Next we address the partial factor correlation between a test and an external variable, conditional on test anxiety. The penultimate section contains the example. We conclude by discussing the significance of these results for research on test anxiety and for applications of CFA that face similar types of rotational indeterminacy.

**A CFA Model of Test Anxiety**

Test anxiety is again receiving more attention in practical settings, especially with relevance for high-stakes testing in the school context (e.g., Cizek & Burg, 2006; van der Embse & Hasson, 2012; Meijer & Oostdam, 2007; Parks-Stamm, Gollwitzer, & Oettingen, 2007; Putwain, 2007; Putwain & Daniels, 2010; Weems et al., 2010), and in the personnel selection context (e.g., Dobson, 2000; McCarthy & Goffin, 2005; Proost, Derous, Schreurs, Hagtvet, & Witte, 2008). Test anxiety also has theoretical importance, for instance when the relationship between personality and intelligence is at stake (e.g., Furham, Forde, & Cotter, 1998; Moutafi, Furnham, & Tsaousis, 2006; Zeidner, 1995), and in explaining group differences in test performance (e.g., Araki, 1992; Maria & Nuovo, 1990; Zeidner & Safir, 1989). In this section we provide a selective review of this literature with the purpose of developing a general model that incorporates the substantive theory. We first discuss the role of test anxiety in the measurement of ability and then the relationships among test anxiety, test scores, and external variables.

*Test Anxiety and the Measurement of Ability*

It is well established that test anxiety is negatively correlated with aptitude test scores (Ackerman & Heggestad, 1997; Hembree, 1988; Zeidner, 1998). Oostdam and Meijer (2003) considered two main hypotheses to explain this correlation. The *deficit hypothesis* states that test takers with a skill deficit (i.e., lower ability) experience
higher anxiety. In its strongest phrasing, the hypothesized skill deficit fully explains the observed correlation between test scores and test anxiety. Several authors have argued for a variation of this hypothesis (e.g., Ball, 1995; Pekrun, 1992; Tobias, 1992).

In terms of CFA, we interpret the deficit hypothesis to mean that ability and test anxiety are correlated. This is depicted in Figure 1 by the factor correlation ($\phi_{AX}$). The circles labeled $\theta_X$ and $\theta_A$ denote ability and test anxiety, respectively. The squares labeled $X$ and $A$ denote their indicator variable(s). We do not specify whether ability causes test anxiety or vise versa, nor do we eliminate the possibility that they are related through some third variable. Such considerations are of central importance in the test anxiety literature (see McDonald, 2001), but they take us beyond the scope of measuring ability.

**Figure 1 Here**

The second explanation considered by Oostdam and Meijer (2003) is the *interference hypothesis*. Under this hypothesis, test anxiety prevents anxious test takers from performing at their true level of ability. They provided support for this hypothesis by showing that students with higher test anxiety perform better in less stressful situations. Several other studies have also provided evidence for the interference hypothesis. Hembree (1988) listed a large number of anxiety reducing interventions that yield improvements in performance for anxious test takers. Parks-Stamm, Gollwitzer, and Oettingen (2007) found an interaction between test anxiety and test taking strategy. Students with high test anxiety perform better with a temptation inhibiting strategy (staying away from distractions) than with a task facilitating strategy (intensifying one’s efforts on the test), whereas the effect is opposite and much smaller for students with low test anxiety. Their result again shows that the relationship between test scores and test anxiety can be manipulated independently of ability, which supports the interference hypothesis.

In Figure 1 we represent the interference hypothesis by letting test scores load on test anxiety ($\lambda_{AX}$). This means that test scores are not conditionally independent given ability, or equivalently, that the test is not a reliable measure of ability. The
presence of a nonzero cross-loading has serious implications for educational practice (Cizek & Burg, 2006).

These considerations lead us to a two-dimensional model of test scores in which the factors, ability and test anxiety, are correlated. A similar model was arrived at by Wicherts and Zand Scholten (2010). The model has the following advantages. Firstly, it disentangles the contributions of ability and test anxiety to test performance. These contributions are represented by the factor loadings $\lambda_X$ and $\lambda_{AX}$, respectively. Secondly, even though test scores ($X$) may be confounded by test anxiety ($\theta_Y$), the resulting measure of ability ($\theta_X$) is not. In other words, we obtain a reliable measure of ability for anxious test takers. Additionally, the model allows us to test hypotheses about test anxiety within a CFA framework.

Unfortunately, the model in Figure 1 is subject to rotational indeterminacy. While this will be obvious to the experienced factor analyst, it is perhaps less obvious that the usual identification constraints employed in CFA have non-trivial implications for the effects of interest. Identification of the model is discussed in the following section.

The model presented in Figure 1 does not exhaust the possible explanations of the correlation between test scores and test anxiety. For example, a direct effect of $X$ on $A$ was described by Deffenbacher (1978). This effect was induced by a particular experimental situation that is not reflective of usual testing practices, so we do not address it here. Zeidner (1995) discussed the possibility that measures of test anxiety may be confounded by cognitive ability, with less able students perceiving themselves as having higher anxiety. We are not aware of any evidence to support this possibility and therefore do not consider it further. We also note that both ability and test anxiety can be treated as multidimensional. Several authors have discussed this kind of approach to test anxiety (e.g., Meijer & Oostdam, 2007; Wren & Benson, 2004), and multidimensional models of ability are well known (e.g., Carroll, 1993). It is then possible to include multiple latent variables for both ability and test anxiety, with a structure similar to that in Figure 1 being possible for each pairwise matching of the two types of latent variables. Despite these limitations, our view is that any serious
attempt to model the role of test anxiety in the measurement of ability must minimally include the effects depicted in Figure 1.

*Test Anxiety, Test Scores, and External Variables*

The deficit and interference effects described above also have consequences for correlation studies involving aptitude tests. In the context of validity studies, evident external variables are school grades (e.g., Hembree, 1988; Pintrich & Groot, 1990) and job performance (e.g. Bertua, Anderson, & Salgado, 2005; Hunter & Hunter, 1984; Salgado & Anderson, 2003). The relationship of test anxiety with school grades is negative, while the direction of the relationship with job performance is unclear. Both school grades and job performance are often evaluated in a testing context, in which case the foregoing considerations about measurement are again relevant.

**Figure 2 Here**

In order to capture the relationship between a test and a single external variable ($\theta_Y$), we extend the model in Figure 1 as shown in Figure 2. The correlation between ability and the external variable is denoted $\phi_{XY}$. The paths labelled $\phi_{AY}$ and $\lambda_{AY}$ are the deficit and interference effects of test anxiety on the external variable, respectively. These may or may not be applicable in a given context. We refer to the submodel corresponding to Figure 1 as the AT (anxiety – test) model. The corresponding effects running between test anxiety and the external variable we refer to as the AE (anxiety – external variable) model. The complete model is comprised by the AT submodel, the AE submodel, and the correlation between $\theta_X$ and $\theta_Y$.

Although it is common practice to use the manifest correlation between $X$ and $Y$ to measure the association between ability and external variables, this can hardly be recommended. Wicherts and Zand Scholten (2010) have shown that manifest correlations are biased in the presence of interference effects. It is therefore preferable to take a latent variable approach. Nonetheless, we show that the partial correlation between $X$ and $Y$, controlling for $A$, is not affected by interference effects when using usual the univariate correction for attenuation. This result is useful for researchers who are committed to using observed correlations.
Summary

This section has selectively reviewed the literature on test anxiety in order to develop an appropriate CFA model. The model requires that test scores, and possibly external variables, are treated as two-dimensional with correlated factors. In the following section we discuss the identification of this model.

Confound Indeterminacy in the AT Model

This section focuses exclusively on the AT model depicted in Figure 1. Entirely similar considerations come into play when addressing the AE model, so that we can generalize to the full model depicted in Figure 2 without needing to explicitly consider both submodels. We begin by showing that the AT model is not identified, and in particular that it is subject to a specific type of rotational indeterminacy, which we refer to as confound indeterminacy. Following the usual practice in CFA, the model would be identified by imposing a simple structure on a subset of the indicator variables and/or requiring an orthogonal factor solution (see Millsap, 2001). We show that these constraints render the effects of interest uninterpretable. We therefore consider alternative methods of identifying the model.

As noted, the identification of the AT model is related to the topic of rotational indeterminacy. For reference we provide a brief review of rotational indeterminacy here. The basic idea is that, in the unrestricted multidimensional factor model, factor scores and factor loadings can be linearly transformed without changing the model-implied covariance matrix (see Millsap, 2001; Yanai & Ichikawa, 2007). Therefore rotational indeterminacy is a type of model unidentification – numerically different parameter arrangements correspond to numerically identical observed moments, and hence numerically identical goodness of fit. In order to deal with rotational indeterminacy, some of the model parameters must be set to fixed values. In exploratory factor analysis this is accomplished by means of various analytic factor rotations (see Browne, 2001, for a review). In CFA, structural properties are imposed on the factor pattern and/or the factor covariance matrix until sufficient constraints have been given to eliminate rotational indeterminacy. As we now discuss, the factor
pattern implied by the AT model is not sufficiently constrained to eliminate rotational indeterminacy.

Rotational Indeterminacy in the AT Model

There are numerous ways to establish the (un)-identification of models for covariance structures (e.g., Bekker, Merckens, & Wansbeek, 1989; Bollen, 1989; Bollen & Bauldry, 2010). The present case allows for a simple algebraic approach. We show that the AT model is subject to rotational indeterminacy by providing a general form for the rotation matrix. This has the advantage of allowing us to explicitly consider what restrictions will serve to identify the model and to describe the consequences of those restrictions for the effects of interest. We assume throughout that the scales of the latent variables are fixed by setting their means to zero and their variances to one. As is usual, we treat rescaling of the latent variables as a trivial (i.e., ignorable) form of rotational indeterminacy.

The matrix of factor loadings corresponding to Figure 1 is

\[ \Lambda = \begin{bmatrix} \lambda_{X_1} & \lambda_{AX_1} \\ \lambda_{X_2} & \lambda_{AX_2} \\ \vdots & \vdots \\ \lambda_{X_J} & \lambda_{AX_J} \\ 0 & \lambda_{A_1} \\ 0 & \lambda_{A_2} \\ \vdots & \vdots \\ 0 & \lambda_{A_K} \end{bmatrix} \]  

(1)

with \( j = 1, \ldots, J \) indexing the indicators of \( \theta_X \) and \( k = 1, \ldots, K \) indexing the indicators of \( \theta_A \). The loadings \( \lambda_{X_j} \) and \( \lambda_{A_k} \) correspond to the measurement models of \( \theta_X \) and \( \theta_A \) respectively, and we assume that these are non-zero. The cross loadings \( \lambda_{AX_j} \) denote the interference effects of \( \theta_A \) on the \( X_j \). The zeros are a structural restriction given by the assumption that indicators of test anxiety do not load on ability. In keeping with Figure 1, the factor correlation is a free parameter that
represents the deficit effect. In this section we denote the factor correlation as $\phi$.

Rotational indeterminacy of $\Lambda$ can be demonstrated by showing the existence of a nonsingular rotation matrix $R$ such that $\Lambda R$ contains the same structural zeros as $\Lambda$. It can be verified that any such rotation matrix has the form

$$R = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \quad (2)$$

where $a, b, c$ are real numbers with $a \neq 0 \neq c$ determining the scale of the factors.

Setting $b = 0$ identifies the model up to the scale of the factors. This was established using linear algebra package of Maple 10 (Maple Version 10 [computer software], 2005) to evaluate the Wald rank rule for the transformed model (see Bollen & Bauldry, 2010). Therefore rotational indeterminacy is the only source of unidentification in the AT model, and non-trivial rotational indeterminacy occurs only when $b \neq 0$. We refer to this type of indeterminacy as confound indeterminacy.

In order to eliminate confound indeterminacy, we require additional constraints on the AT model that are violated whenever $b \neq 0$. The possible constraints can be determined by computing the rotation. Computation of $\Lambda R$ shows that $b$ appears in the cross-loadings, $\lambda_{AX_j}$, but not the other factor loadings. The rotated values are

$$\lambda^*_{AX_j} = b \lambda_{X_j} + c \lambda_{AX_j} \quad (3)$$

for $j = 1, \cdots, J$. The rotated factor correlation can be written

$$\phi^* = \phi - \frac{b/c}{\sqrt{(b/c)^2 - 2\phi b/c + 1}}. \quad (4)$$

Equations (3) and (4) show that, in order to eliminate confound indeterminacy, it is minimally required that one of the interference effects or the deficit effect be fixed. For example, if we require that $\lambda^*_{AX_j} = c \lambda_{AX_j} = r$, for fixed values of $j$ and $r \in \mathbb{R}$, this requirement can only be satisfied if $b = 0$. Therefore, fixing one of the interference effects serves to identify the model. Similarly, requiring $\phi^* = \phi = s$ for a fixed value of $s \in (-1, 1)$ also identifies the model.
In the usual practice of CFA, we would identify the AT model by setting either \( r \) or \( s \) to an arbitrary numerical value, and the value chosen is most commonly zero. Setting \( r = 0 \) is equivalent to forcing a simple structure on a subset of the factor pattern; setting \( s = 0 \) is equivalent to requiring an orthogonal factor solution. Thus it appears that we must “give up” one of the effects of interest in order to identify the model. However, the situation is in fact far worse than this. As we show below, selecting an arbitrary value for the deficit effect also implies that the numerical values of the interference effects are arbitrary. A similar situation holds when fixing one of the interference effects. In short, if we identify the AT model using the usual parameter constraints employed in CFA, \textit{none} of the effects of interest are interpretable.

\textit{Consequences of Usual CFA Identification Constraints for the AT Model}

In the following we describe the consequences of identification constraints placed on the AT model in terms of factor rotations applied to an unknown data-generating model. Model parameters without superscripts (e.g., \( \phi, \lambda_{X_j} \)) denote the data-generating parameters. These can be thought of as the true parameters in the population of interest. Identification constraints are described in terms of factor rotations applied to the data-generating parameters. Quantities relating to rotations are denoted by latin letters (e.g., \( R, r \)) and are due to identification choices made by researcher. Model parameters with the asterisk superscript (e.g., \( \phi^*, \lambda_{X_j}^* \)) denote the parameter values of an identified model. These are regarded as functions of the true parameter values and the chosen identification constraints.

Consider first the implications of fixing the interference effect on \( X_j \) to some arbitrary value \( r \in \mathbb{R} \), where \( j \) and \( r \) are according to the convenience of the researcher. This can be interpreted as choosing a rotation matrix \( R \) such that \( \lambda_{AX_j}^* = r \). Inserting this restriction into equation (3) implies that

\[
b = \left( r - c \lambda_{AX_j} \right) / \lambda_{X_j}.
\]

Note that setting \( r = 0 \), which would be usual, does not imply that \( b = 0 \). Thus the identification constraint represents a non-trivial rotation of the data-generating
parameters. In order to have $b = 0$, we must have $r = c\lambda_{AX_j}$. This means that the researcher would have to guess the true interference effect in order to identify the model in this manner.

We can compute the value of the rotated factor correlation implied by equation (5) using equation (4):

$$
\phi^* = \frac{\phi - (r/c - \lambda_{AX_j})/\lambda_{X_j}}{\sqrt{\phi - (r/c - \lambda_{AX_j})/\lambda_{X_j})^2 - \phi^2 + 1}}.
$$

It is complicated function involving both the value chosen for the constraint, $r$, and the value of the data-generating parameter, $\lambda_{AX_j}$. Figure 3 shows equation (6) for $\phi = 0$, $r = 0$, and various combinations of $\lambda_{AX_j}$ and $\lambda_{X_j}$. This corresponds to a case in which there is no deficit effect in the population ($\phi = 0$), and the model is identified by setting one of the measurement interference effects to zero ($r = 0$). The figure shows how the model-implied deficit effect, $\phi^*$, departs from the true value of zero. Specifically, it is seen that $\phi^*$ is non-linear function of the true measurement interference effect, $\lambda_{AX_j}$, and that this relationship is moderated by $\lambda_{X_j}$.

In general, identifying the AT model by fixing one of the interference effects to an arbitrary value will result in a model-implied deficit effect that is not equal to the deficit effect in the population of interest. In research applications, this means that the numerical value of the estimated deficit effect is uninterpretable. In particular, testing hypotheses about the value of $\phi^*$ does not allow one to make conclusions about $\phi$.

**Figure 4**

Fixing one of the interference effects also has consequences for the remaining interference effects. Substituting equation (5) into equation (3), we have

$$
\lambda_{AX_k}^* = c\lambda_{AX_k} - (r - c\lambda_{AX_j}\lambda_{A_k})/\lambda_{X_j} \quad j \neq k.
$$

Therefore, when one interference effect is fixed, the remaining interference effects are a function of the value to which the effect is fixed, $r$, and the value of the data-generating parameter, $\lambda_{AX_j}$, as well as other quantities. The functional
relationship here is similar to that considered directly, so we do not discuss it further.

We now address the case where the model is identified by fixing the value of the factor correlation. This is equivalent to choosing a rotation such that $\phi^* = s$, for some arbitrary fixed value $s \in (-1, 1)$. Then equation (4) implies

$$b/c = \sqrt{s^2(1 - \phi^2)} - \phi$$

and substitution of equation (8) into equation (3) shows that

$$\lambda_{AX_j}^*/c = \lambda_{AX_j} + \left( \sqrt{s^2(1 - \phi^2)} - \phi \right) \lambda_{X_j}$$

for $j = 1, \ldots, J$. It can be verified $\lambda_{AX_j}^*/c = \lambda_{AX_j}$ only when $\phi = s$. Thus we arrive at a situation similar to that described above: identifying the AT model by fixing the deficit effect to an arbitrary value results in model-implied interference effects that are not, in general, equal to the interference effects in the population of interest.

The right hand panel of Figure 4 plots equation (9) for $\lambda_{AX_j} = 0$, $s = 0$, and $c = 1$. This corresponds to a case in which there is no interference effect on variable $X_j$ in the population of interest and the model is identified by an orthogonal factor solution. The model-implied interference effect, $\lambda_{AX_j}^*$, is seen to be a linear function of the true deficit effect, $\phi$, and this is again moderated by the reliability of $X_j$.

Summary. The results of this subsection can be conveniently summarized in terms of a trade-off between the interference effects and the deficit effect: identification of the model requires fixing one, but doing so influences the numerical value of the other. Consequently, none of the effects of interest can be interpreted if the model is identified by arbitrary parameter constraints. In applications, this means that the numerical values of the parameter estimates and their related significance tests would be strictly meaningless. Therefore we turn to consider alternative means of identifying the AT model.
Test Anxiety 15

Testable Identification Constraints on the AT Model

The foregoing has made it clear that “non-arbitrary” identification constraints are required in order to interpret the AT model. The strategy we pursue here is to introduce more constraints than are necessary for identification, and in particular we consider equality restrictions among the interference effects. These constraints can be seen as simplifying assumptions or can be motivated by substantive hypotheses about test anxiety. The resulting models are properly nested within the just-identified AT model, so the validity of the constraints can be tested via the usual procedures for nested models (e.g., Bollen, 1989; Satorra & Bentler, 2001). Thus we replace the problem of identification with that of testing hypotheses about the interference effects.

The proposed solution does come with disadvantages. Although the constraints we consider serve to formally identify the model, there exist cases where the model may be numerically unidentified due to specific configurations of the ability loadings. We describe these configurations below. Additionally, it may turn out that none of the available constraints are compatible with a given data set (i.e., the nested hypotheses are all rejected). Although it is not ideal, our approach to identification does allow test anxiety to be modelled within a CFA framework; this is a distinct advantage over the use of arbitrary identification constraints.

In terms of the previous section, equality constraints on the interference effects require setting $\lambda_{AX_j} = \lambda_{AX_k}$ for some choices of $j \neq k$. Substituting this requirement into equation (5) shows that

$$b = \frac{(\lambda_{AX_k} - c\lambda_{AX_j})}{\lambda_{A_j}}.$$  \hspace{1cm} (10)

Assuming that $c = 1$, then $b = 0$ just in case $\lambda_{AX_j} = \lambda_{AX_k}$. In other words, if it is true that $\lambda_{AX_j} = \lambda_{AX_k}$ in the population of interest, then using this equality constraint to identify the model implies that $b = 0$, and therefore that we can interpret the effects of interest. Accordingly, we require a test of the hypothesis that $\lambda_{AX_j} = \lambda_{AX_k}$.

A single equality constraint on the interference effects is only sufficient to just-identify the AT model. Therefore we can not test the single constraint. Imposing
multiple equality constraints across the interference effects, however, does yield a testable hypothesis in the form of a properly nested model. For example, we may hypothesize that the interference effect is constant over all $j = 1, \ldots, J$ ability indicators, resulting in a total of $J - 1$ parameter restrictions on the AT model. This may be referred to as a tau-equivalent model for the interference effects. To obtain the nesting model, the AT model can be just-identified by imposing any one of the constraints discussed in the foregoing. Assuming that the AT model contains the data-generating process, the likelihood ratio statistic has an asymptotic chi-square distribution on $J - 2$ degrees of freedom under the null hypothesis.

In CFA it is usual to impose tau-equivalence as a numerical simplification, without having a specific empirical motivation outside of the fact that the data is congenial to the simplification. In the context of test anxiety, tau-equivalence of the interference effects may also represent a substantive hypothesis about test anxiety. For example, if the ability indicators are different test items on a single test, tau-equivalence means that each item is equally anxiety provoking, regardless of its content or its position within the test. Such a hypothesis was addressed by Oostdam and Meijer (2003). If the ability indicators are different testing situations, tau-equivalence means that test takers experience equal amounts of test anxiety regardless of the situation. On the other hand, it is commonly supposed that test anxiety should be increased in high-stakes testing as compared to low-stakes testing (e.g., Cizek & Burg, 2006). This supposition can also be tested under the AT model. For example, we could have test takers write a series of examinations (or exam items) in either a high-stakes or low-stakes situation and test whether tau-equivalence holds within but not between the two situations. Further equality constraints can be adduced from the test anxiety literature. We provide an illustration of this approach to model identification in our example.

Before moving on we point out an important limitation. Consider the following
two interference effects from equation (3) obtained by using the equality $\lambda_{AX_j} = \lambda_{AX_k}$:

\[
\begin{align*}
\lambda^*_{AX_j} &= b\lambda_{X_j} + c\lambda_{AX_k} \quad j \neq k \quad \text{and} \\
\lambda^*_{AX_k} &= b\lambda_{X_k} + c\lambda_{AX_k}.
\end{align*}
\]

(11)

Equations (11) show that there are in fact two cases in which the rotated factor loadings retain the equality constraint (i.e., $\lambda^*_{AX_j} = \lambda^*_{AX_k}$). The case we have described above occurs when $b = 0$. The second case is when $\lambda_{X_j} = \lambda_{X_k}$, that is, when tau-equivalence also holds for the ability indicators. In this case, the equality restriction in equation (10) does not identify the model. In practice, this is recognizable by non-convergence of estimation algorithms and by obtaining different estimates from disparate starting values. In some situations this can be remedied by different choices of $j$ and/or $k$. However, if a tau-equivalent model holds for the ability indicators, the identification strategy proposed here will no longer work.

**Summary**

The general message to be taken from this section is that the usual parameter constraints employed in CFA are not tenable for modelling test anxiety. To deal with this problem we have discussed some non-trivial (i.e., testable) parameter constraints that serve to identify the AT model. These constraints readily correspond to substantive hypotheses about test anxiety, and therefore provide an appropriate solution to the identification problem in the present context.

As noted at the outset of this section, entirely similar considerations apply to the AE model shown in the right hand side of Figure 2. Therefore two sets of constraints of the kind discussed here must be imposed in order to identify the full model. Given these constraints, the correlation between ability and an external variable can be estimated, in addition to the deficit and interference effects. We illustrate this in our example.

As noted, the constraints we have discussed are not effective when the ability indicators are tau-equivalent. Moreover, full-blown structural equation modelling is not a viable option in many correlation studies. In such cases it is desirable to have a
alternative measure of association for test scores that is not confounded by test anxiety. This is the topic of the following section.

**A Structural Relation That is Invariant Under Confound Indeterminacy**

In this section we return to the full model depicted in Figure 2. We show that the partial correlation between $\theta_X$ and an external variable $\theta_Y$, given $\theta_X$, is invariant under confound indeterminacy. This means that the partial correlation provides a measure of association between $\theta_X$ and $\theta_Y$ that is interpretable regardless of the identification constraints employed. We also derive the partial correlation between $X$ and $Y$. This is a biased estimator of the latent partial correlation, but the usual correction for attenuation serves to correct this bias. Thus it is possible to estimate the partial correlation between $\theta_X$ and $\theta_Y$ without estimating the CFA model.

From a methodological perspective, this is a curious result that presents interesting possibilities for generalisation. From the perspective of correlational research involving aptitude tests, it means that we can control for the confounding effects of test anxiety without having to directly model those effects. Of course, the applicability of the partial correlation will depend on the research question at hand.

Let

$$
\begin{align*}
\mathbf{f} &= \begin{bmatrix} \theta_X \\ \theta_Y \\ \theta_A \end{bmatrix} \\
\Phi &= \begin{bmatrix} 1 & - & - \\ \phi_{XY} & 1 & - \\ \phi_{AX} & \phi_{AY} & 1 \end{bmatrix} \\
\mathbf{L} &= \begin{bmatrix} \lambda_X & 0 & \lambda_{AX} \\ 0 & \lambda_Y & \lambda_{AY} \\ 0 & 0 & \lambda_A \end{bmatrix}
\end{align*}
$$

describe the full model in Figure 2. The vector $\mathbf{f}$ contains the factors and $\Phi$ is their correlation matrix. $\mathbf{L}$ gives an abbreviated factor pattern. Each row of the complete factor pattern must have structural zeros corresponding to one of the three rows shown in $\mathbf{L}$. Again we assume that the scales of the factors are fixed by setting their means to zero and variances to one.

As with the AT and AE submodels, the full model is subject to confound indeterminacy. The rotation matrix for the factor pattern, $\mathbf{T}$, and the inverse rotation
that transforms the factors, $T^{-1}$, have the form

$$T = \begin{bmatrix} t_X & 0 & t_{AX} \\ 0 & t_Y & t_{AY} \\ 0 & 0 & t_A \end{bmatrix}$$

$$T^{-1} = \begin{bmatrix} \frac{1}{t_X} & 0 & -\frac{t_{AX}}{t_X t_A} \\ 0 & \frac{1}{t_Y} & -\frac{t_{AY}}{t_Y t_A} \\ 0 & 0 & \frac{1}{t_A} \end{bmatrix}$$  \hspace{1cm} (13)$$

The null entries are implied by the structural zeros in the factor pattern, as is readily verified. The other values are assumed to be non-zero real numbers. For later reference we note that $T$, $T^{-1}$, $L$ and $L^{-1}$ all have the same structural zeros. This means that the abbreviated factor pattern and its inverse are admissible rotations under confound indeterminacy.

Let $\phi_{XY|A}$ denote the partial correlation between $\theta_X$ and $\theta_Y$ conditional on $\theta_A$. Also let $\phi_{XY|A}^*$ denote the partial correlation resulting from application of the rotation $T$ to the model. In the appendix we show that

$$\phi_{XY|A}^* = \phi_{XY|A}. \hspace{1cm} (14)$$

Equation (14) says that the partial correlation of $\theta_X$ and $\theta_Y$, given $\theta_A$, is invariant under confound indeterminacy. Therefore the value of $\phi_{XY|A}$ is not affected by the constraints used to identify the model. The following section provides an empirical illustration of this fact. We note that the other partial correlations that can be computed between the factors are affected by confound indeterminacy (see appendix).

We now consider the manifest partial correlation between an aptitude test $X$ and an external variable $Y$ when controlling for an anxiety indicator $A$. The zero-order covariance matrix of the three indicators can be written as $\Sigma = LL' + \Psi$, where $L$ and $\Phi$ are from equation (12) and $\Psi$ is the diagonal matrix of error variances. Let $\Sigma^* = \Sigma - \Psi = LL'$ and let $\rho_{XY|A}^*$ denote the partial correlation between $X$ and $Y$ computed on $\Sigma^*$. As noted in connection with equation (13), $L^{-1}$ is in fact an admissible rotation under confound indeterminacy. Thus we can write $T = L^{-1}$ so that $\Sigma^* = T^{-1} \Phi T^{-1}$ is
seen to be a rotation of the factor correlation matrix. Then equation (14) implies

$$\rho_{YX|A}^* = \phi_{XY|A}.$$  \hspace{1cm} (15)

From equation (15) it is clear that the manifest partial correlation computed on $\Sigma$ is a biased estimator of the latent partial correlation, and that this bias is due only to the measurement error $\Psi$. Therefore, under an appropriate correction for attenuation, the manifest partial correlation becomes an unbiased estimate the latent partial correlation.

It is readily shown that the usual correction for attenuation for the partial correlation (Stouffer, 1936) is also applicable in the current situation. This is both surprising and convenient. It is surprising because the correction for the partial correlation is motivated through the true score theory, and in the present case $X$ and $Y$ are not compatible with true score theory due their multidimensionality. It is convenient because it means that only the three indicators and their reliabilities are required in order to obtain an unbiased estimate of the latent partial correlation. If the indicators have known reliability coefficients, the sample size requirements are equivalent to those for estimating the zero-order correlations (e.g., $N \geq 30$). Thus the result is suitable for application in correlation studies. We illustrate the use of the attenuation correction for the partial correlation in the following section.

**Summary**

This section has shown that the partial correlation between $\theta_X$ and an external variable $\theta_Y$ conditional on $\theta_A$ is invariant under confound indeterminacy. Moreover, this correlation can be estimated using the usual correction for attenuation for the manifest partial correlation. This result allows researchers to control for the confounding effects of test anxiety in correlation studies involving aptitude tests.

**Empirical Example**

This section describes an example from educational testing. The purpose of the example is to illustrate (a) the consequences of the usual CFA model identification
constraints on the effects of interest, (b) the use of tau-equivalent interference effects to identify the model, and (c) estimation of the factor correlation and partial correlation. We emphasize at the outset that we pre-selected indicator variables in order to suit this purpose, and therefore the example is only illustrative and does not provide a basis for substantive conclusions.

The data were obtained from voluntary pre-testing of 898 high school students registered for the 2009 Vestibular examinations at the University of Brasilia. Prior to sitting four aptitude tests (Humanities, Linguistics, Math, and Natural Science), students responded to a questionnaire containing 39 five-point Likert items about cognitive, emotional, and physical aspects of test anxiety. The aptitude tests were administered according to an incomplete block design in which each test consisted of four blocks of 15 items and each student responded to three blocks from each test. The full data set and the test anxiety items are described in detail by da-Silva and Gomes (2011).

Our analyses were necessarily performed at the item level. We employed the weighted least square estimator as implemented in Mplus 5.21 (Mplus 5.21 [computer software], 2009). This estimator is explicitly related to the foregoing discussion of continuous indicators via the polychoric correlation matrix (see e.g., Muthén, 1978). Our main conclusions were not affected by use of other estimators (e.g., marginal maximum likelihood).

The present example made use of $N = 382$ complete responses to nine items from both the Linguistics (L) and Natural Sciences (NS) examinations. We also made use of items 7, 26, 30, 31, and 34 of the test anxiety questionnaire, which were all related to the experience of distracting thoughts (e.g., “I feel upset because of thoughts that distract me”). Items were pre-selected based on their unidimensional measurement models and the correlations among their sumscores.

The distraction items showed a stronger sumscore correlation with test performance than the other test anxiety items, which is in keeping with previous research (e.g., Oostdam & Meijer, 2003; Parks-Stamm et al., 2007). However, none of the correlations were very large ($-0.2 < r < 0$). Examination of the data showed that
most students reported low test anxiety regardless of their subsequent performance. We interpret this to mean that students perceived the voluntary pre-testing context as low-stakes. Nonetheless, the sumscore correlations for the selected items were reasonable from the perspective of the literature on test anxiety (Reeve et al., 2009). The correlations are reported in Table 1.

Identifying the Model

We fitted the full model in Figure 2 with three different identification constraints. The results are summarized in Table 2. The first two models are just-identified, while the third is properly nested within the just-identified model. As reported in the second and third columns, the overall fit of the models was marginal but sufficient for the purpose of illustration.

In Model 1 the deficit effects for the L and NS factors were set to zero. The table reports the strongest interference effects for two sets of items. Both effects are negative and significant at the .01 level, which is in keeping with the interference hypothesis. The zero-order and partial correlations between the L and NS factors are equal, which follows from the identification constraints.

In Model 2 we fixed two of the interference effects to zero instead. In particular, we fixed the L and NS items for which the ratio of the interference effect to the ability loading was largest in Model 1. As described in Figure 3, this should lead to a solution that is numerically disparate from Model 1. This conclusion is borne out in Table 2. It is seen that there is a now sizeable deficit effect for both L and NS items, but the largest “interference effects” are in the wrong direction (i.e., more test anxiety leads to better performance on these items). Also note that while the zero-order factor
correlation is larger in Model 2 than Model 1, the partial factor correlation is constant. This illustrates the result proven in equation (14) above.

Model 3 was obtained by forcing subsets of the inference effects to be tau-equivalent. We did not have any substantive hypotheses about the interference effects, so the model is merely exploratory. Setting tau-equivalence over all items, or only within the L and NS items, did not result in acceptable fit. Instead we found four sets of equivalent items. The first set contained all items from both tests that did not show interference. The second contained the remaining L items, whose interference effect is reported in Table 2. The third set contained five NS items with a moderate interference effect; this was substantially larger than that of the L items. The final set contained a single NS item that showed a very strong interference effect, which is reported in the table. The chi-square difference test against the just-identified model showed the constraints to be acceptable ($\chi^2 (df = 13) = 13.33, p = .423$). The fit statistics reported in the table corroborate this conclusion. Neither of the resulting deficit effects were significant at the .01 level, and the factor correlations were similar those of Model 1.

In summary, comparison of Models 1 and 2 indicates the arbitrariness of the direction, numerical magnitude, and statistical significance of the effects of interest in the just-identified model, and in particular for cases where the usual CFA constraints are employed. Model 3 shows that it is possible to identify the model using equalities over the interference effects, and that the acceptability of these equalities can be tested against the data. In the present example, the resulting model showed no deficit effects but moderate to large interference effects over some but not all test items. This means that test anxiety was associated with worse test performance on some items, but it was not correlated with ability. The discovery of such items has serious implications for educational practice.

Measuring the Association Between Language and Natural Science

We now consider the relationship between the L and NS items in more detail. Table 3 shows various ways of estimating this relationship. The latent correlations are those from Model 3, which we take as the best available estimates. The sumscore
correlations are computed in the usual manner, and the corrected correlations used
coefficient alpha to correct for attenuation in the sumscore correlations. Sumscores
were computed using only the L and NS items that showed tau equivalence in both
their ability loadings and their inference effects in Model 3. This precaution was taken
to ensure that the sumscore is an appropriate summary statistic.

**Table 3 Here**

Referring to Table 3, the sumscore correlations are seen to severely
underestimate the latent correlations. Bias of the sumscore correlations is to be
expected in the presence of interference effects (e.g., Wicherts and Zand Scholten,
2010) and in our case attenuation also plays an important role since the reliability of
the sumscores was not high ($\alpha < .6$). The discrepancy between the sumscore
correlations and the latent correlations illustrates the usefulness of the approach taken
in this paper. The corrected correlations do better, and in particular the corrected
partial correlation is an acceptable approximation of the latent partial correlation.
This illustrates the result given in equation (15).

**Conclusions**

In this paper we have used CFA to model the role of test anxiety in the
measurement of ability and in correlation studies involving aptitude tests. Our main
conclusions are as follows. (a) In the context of test anxiety, we are faced with a
particular type of rotational indeterminacy that we have referred to as confound
indeterminacy. (b) The usual parameter constraints employed for model identification
in CFA are not a viable option for eliminating confound indeterminacy, because they
render the effects of interest uninterpretable. (c) Fortunately, equality constraints
among the cross-loadings of the ability indicators on test anxiety serve to identify the
model and these constraints can be statistically tested. (d) If such constraints are not
tenable for a given data set, we may nevertheless estimate the partial factor correlation
between ability and an external variable, while controlling for test anxiety, because
this correlation remains invariant under confound indeterminacy. (e) Moreover, the partial factor correlation can be estimated using the usual correction for attenuation, which is useful for researchers who are committed to using observed correlations.

With regard to test anxiety research, the model presented here allows us to disentangle the contributions of ability and test anxiety to test performance. It also allows for unbiased estimation of the correlation between ability and other variables of interest, and can be used to test various substantive hypotheses. With regard to testing practice, application of the model allows for identification of tests/item that are biased against anxious test takers, and at the same time provides a measure of ability that is not confounded by test anxiety. In other words, the model yields reliable ability estimates even for test takers who experience anxiety. In general, the CFA approach makes available a large number of useful results for applications concerning test anxiety.

The proposed model has also motivated some methodological considerations. Although we have only discussed these within the context of test anxiety, we hope that further research will generalize our results. Specifically, it would be nice to have a rigorous definition of confound indeterminacy for an arbitrary number of factors, as well as the parameter constraints that serve to eliminate it. It would also be nice to know which partial correlations are invariant under this indeterminacy. Additionally, while we have focused on the consequences of confound indeterminacy for parameter estimates, it is also important to describe the consequences for their standard errors. Because confound indeterminacy is arguably quite pervasive in social science applications, these results could have broad consequences.

Appendix

In this appendix we use equations (12) and (13) to prove equation (14). We begin by computing the transformed factor covariance matrix $R = T^{-1} \Phi T^{-1'}$: 
Writing $S = \text{diag}(R)^{-1/2}$, the transformed factor correlation matrix is given by

\[ \Phi^* = SRS. \]

It is clear from inspection of equation (16) that the zero-order correlation between $\theta_X$ and $\theta_Y$ is affected by the rotation. However, it is perhaps surprising to find that this is not the case for the partial correlation between $\theta_X$ and $\theta_Y$ when controlling for $\theta_A$.

The partial correlation between $\theta_X$ and $\theta_Y$ before rotation is defined as

\[ \phi_{XY|A} = \frac{\phi_{XY} - \phi_{AX} \phi_{AY}}{\sqrt{1 - (\phi_{AX})^2} \sqrt{1 - (\phi_{AY})^2}} \]  

The transformed partial correlation, $\phi^*_{XY|A}$, can be obtained by computing it elements directly from $R$. Letting $\sigma^2_{(\cdot)}$ denote the diagonal elements of $R$, we have in the numerator of $\phi^*_{XY|A}$:

\[ \phi^*_{AX} \phi^*_{AY} = \frac{\phi_{AX} \phi_{AY} - \frac{t_{AX}}{t_A} \phi_{AX} - \frac{t_{AY}}{t_A} \phi_{AY} + \frac{t_{AX} t_{AY}}{t_A}}{t_X t_Y \sigma^*_{X} \sigma^*_{Y}} \]

from which it follows that

\[ \phi^*_{XY} - \phi^*_{AX} \phi^*_{AY} = \frac{\phi_{XY} - \phi_{AX} \phi_{AY}}{t_X t_Y \sigma^*_{X} \sigma^*_{Y}}. \]  

For the denominator of $\phi^*_{XY|A}$:

\[ 1 - (\phi^*_{AX})^2 = 1 - \frac{1}{t_X t_A} \left( \frac{\phi_{AX} - \frac{t_{AX}}{t_A}}{\sigma^*_{X} \sigma^*_{A}} \right)^2 = \frac{t^2_X \sigma^2_{X} - \left( \phi_{AX} - \frac{t_{AX}}{t_A} \right)^2}{t^2_X \sigma^2_{X}} = 1 - \phi^2_{AX}. \]
where the last equality follows directly from expansion of the term in parentheses. A similar argument shows that

\[ 1 - (\phi_{AY}^2)^2 = \frac{1 - \phi_{AY}^2}{t_Y^2\sigma_Y^2}. \]  

(21)

From inspection of equations (19), (20), and (21) it is readily seen that \( t_X, t_Y, \sigma_X^*, \) and \( \sigma_Y^* \) cancel out in the numerator and denominator of the transformed partial correlation, leading to equation (14). Following the same sequence of computations, one finds that both of the other partial correlations among \( \theta_X, \theta_Y, \) and \( \theta_A \) are affected by rotation; these do not reduce to convenient expressions.
References


Selection and Assessment, 8, 99–109.


Oostdam, R., & Meijer, J. (2003). Influence of test anxiety on measurement of


Reeve, C. L., & Bonaccio, S. (2008). Does test anxiety induce measurement bias in cognitive ability tests. *Intelligence, 36*.

Reeve, C. L., Heggestad, E. D., & Lievens, F. (2009). Modeling the impact of tests anxiety and test familiarity on the criterion-related validity of cognitive ability tests. *Intelligence, 37*.


Table 1

*Sumscore Correlations for Linguistics, Natural Sciences, and Test Anxiety*

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>NS</th>
<th>ANX</th>
<th>ANX</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS</td>
<td>.406</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANX</td>
<td>-.146</td>
<td>-.194</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note:* L = linguistics; NS = natural science; ANX = test anxiety. All correlations significant at the .01 level.
Table 2  
Model Summary for Three Different Identification Constraints

<table>
<thead>
<tr>
<th>Constraints</th>
<th>RMSEA</th>
<th>CFI</th>
<th>L</th>
<th>NS</th>
<th>L</th>
<th>NS</th>
<th>0-order</th>
<th>partial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>0.066</td>
<td>.928</td>
<td>-.284*</td>
<td>-.307*</td>
<td></td>
<td></td>
<td>.701*</td>
<td>.701*</td>
</tr>
<tr>
<td>Model 2</td>
<td>0.066</td>
<td>.928</td>
<td>.326*</td>
<td>.305*</td>
<td>-.577*</td>
<td>-.523*</td>
<td>.790*</td>
<td>.701*</td>
</tr>
<tr>
<td>Model 3</td>
<td>0.064</td>
<td>.929</td>
<td>-.188*</td>
<td>-.361*</td>
<td>-.118</td>
<td>.102</td>
<td>.669*</td>
<td>.689*</td>
</tr>
</tbody>
</table>

*Note:* RMSEA = root mean square error of approximation; CFI = comparative fit index; L = language items; NS = natural science items. The factor correlation is between L and NS latent variables, with the test anxiety factor held constant for the partial correlation. The reported interference effects are the strongest for each model. ‘*’ denotes a statistical significance at the .01 level. The models are described in the text.
Table 3
*Correlation Coefficients for Linguistics and Natural Sciences Items*

<table>
<thead>
<tr>
<th></th>
<th>Sumscore</th>
<th>Corrected</th>
<th>Latent</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-order</td>
<td>0.270</td>
<td>0.701</td>
<td>0.669</td>
</tr>
<tr>
<td>Partial</td>
<td>0.256</td>
<td>0.686</td>
<td>0.689</td>
</tr>
</tbody>
</table>

*Note:* Sumscore denotes the usual sumscore correlation between language (L) and natural science (NS) items. Corrected denotes the attenuation corrected sumscore correlations. Latent denotes the correlations for Model 3 above and are repeated from Table 2. The partial correlations controlled for test anxiety. The sumscore and corrected correlations were computed using subsets of L and NS items identified as tau-equivalent in Model 3.
**Figure Captions**

*Figure 1.* Two possible explanations of the correlation between test scores ($X$) and measures of test anxiety ($A$). Both $X$ and $A$ may represent multiple indicators. The aptitude or ability being measured is denoted as $\theta_X$ and test anxiety is denoted by $\theta_A$. The correlation $\phi_{AX}$ denotes a deficit effect. The cross-loading $\lambda_{AX}$ denotes an interference effect. The arrows $\lambda_X$ and $\lambda_A$ are the factor loadings of the indicator variables on their measurement constructs. Error terms are suppressed for visual clarity.

*Figure 2.* Extension of the measurement model in Figure 1 to include a third variable of interest, $\theta_Y$, with indicator(s) $Y$ and factor loading(s) $\lambda_Y$. The paths labelled $\phi_{AY}$ and $\lambda_{AY}$ mirror the deficit and interference effects described in Figure 1. The correlation $\phi_{XY}$ represents the relationship between ability and the external variable. The remaining paths are described in Figure 1. Error terms are suppressed for visual clarity.

*Figure 3.* Confounding of null effects by model identification constraints. The left panel shows equation (6) as a function of $\lambda_{AX}$ and $\lambda_X$ with $r = 0$ and $\phi = 0$. Because it is the ratio $\lambda_{AX}/\lambda_X$ that determines the magnitude of $\phi^*$, the raw values of the loadings are not relevant. The right panel shows equation (7) as a function of $\phi$ and $\lambda_X$ with $s = 0$ and $\lambda_{AX} = 0$. Here $c = 1$, so that the magnitude of the factor loadings is interpreted relative the variance of $\theta_A$, which is equal to 1. Since $\phi^*$ and $\lambda^*_{AX}$ are both odd functions, only positive values of the arguments are shown. Values of $\lambda_X$ are given in the right panel.
\[ \phi_{XY} \]

\[ \theta_X \quad \phi_{AX} \quad \phi_{AY} \quad \theta_Y \]

\[ \lambda_X \quad \lambda_{AX} \quad \lambda_A \quad \lambda_{AY} \quad \lambda_Y \]

\[ X \quad A \quad Y \]