Article

Does Acquiescence Affect Individual Items Consistently?

Educational and Psychological Measurement
2015, Vol. 75(5) 764–784

DOI: 10.1177/0013164414560817
epm.sagepub.com
SAGE

Chester Chun Seng Kam[1] and Mingming Zhou[1]

Abstract

Previous research has found the effects of acquiescence to be generally consistent across item *aggregates* within a single survey (i.e., essential tau-equivalence), but it is unknown whether this phenomenon is consistent at the *individual item* level. This article evaluated the often assumed but inadequately tested proposition that individual items are affected by acquiescence to the same degree. We modeled an external acquiescence criterion to assess (a) whether it affected scale items consistently and (b) whether it would be strongly correlated with an acquiescence factor based on an assumption of tau-equivalence. The results did not support this assumption. As further evidence, we identified a situation in which this tau-equivalence assumption could potentially be violated. We propose that the response style be best understood within a framework of an acquiescence × item interaction.

Keywords

response style, acquiescence, bias

There has been a recent surge of interest in studying response styles as an individual difference variable (Bolt, Lu, & Kim, 2014; Rammstedt & Farmer, 2013; Weijters, Geuens, & Schillewaert, 2010b). The term *response style* refers to systematic differences in response scale use between individuals, regardless of item content or respondents' standing on the trait being assessed (Wetzel, Carstensen, & Böhnke, 2013; Wiggins, 1973). A response style that has been shown to occur frequently is acquiescence (disacquiescence), a participant's tendency to overuse one side of a

[1]University of Macau, Taipa, Macau, China

Corresponding Author:
Chester Chun Seng Kam, Faculty of Education, University of Macau, E33, Avenida da Universidade, Taipa, Macau, China.
Email: ChesterKam@umac.mo

scale (e.g., the agreement side; Hui & Triandis, 1985). Recent findings have shown that many response styles, including acquiescence, are stable across measurement scales and time (Billiet & Davidov, 2008; Weijters et al., 2010a, 2010b), implying that acquiescent respondents have a systematic tendency to inflate, or deflate, their item scores. When studying the nature of acquiescence, researchers often make the basic assumption of *(essential) tau-equivalence* as a start, which presumes that the effect of acquiescence is homogeneous across all the measurement items. However, few studies have examined the validity of the tau-equivalence assumption. Based on this assumption, Billiet and McClendon (2000) constructed an acquiescence factor using a structural equation modeling (SEM) technique and found this acquiescence factor to be strongly correlated with another acquiescence measurement. Their results thus supported the tau-equivalence assumption. However, more recent researchers (Weijters et al., 2010b) have contended that the method of measuring acquiescence in Billiet and McClendon's (2000) study was suboptimal, leaving their original conclusion open to question. To deepen our understanding of acquiescence, we investigated the tenability of the tau-equivalence assumption through multiple validation approaches. We also examined the validity of this assumption using multiple scales, to enhance the generalizability of our results.

Nature of Acquiescence

Acquiescence response style (ARS) refers to the tendency of survey participants to agree (or disagree) with an item, even for items with heterogeneous content (Jackson & Messick, 1962). Acquiescence is sometimes called directional bias (Hui & Triandis, 1985) because acquiescent participants may prefer to use one side of a scale, such as ''agree'' and ''strongly agree,'' rather than the other. As a result, ARS is likely to lead to skewness because it can shift the mean of the response distribution to the right or to the left (Schweizer, 2012). ARS can also produce a positive bias in item correlations and weaken negative correlations between regular- and reverse-keyed items, causing a construct to load on two separate factors in factor analysis (Kam & Meyer, 2012; Marsh, 1996). If a survey includes multiple constructs, ARS will strengthen the magnitude of positive correlations and weaken the magnitude of negative correlations. Failure to control acquiescence might thus lead to invalid results (Darcy & Tracey, 2003). Researchers have generally found that the factor structure of their measures improves after the use of a procedure to statistically control for ARS (e.g., Rammstedt & Farmer, 2013; Rammstedt, Kemper, & Borg, 2013). In a SEM model, a better model fit is achieved by specifically modeling the acquiescence factor (e.g., Friborg, Martinussen, & Rosenvinge, 2006).

An important characteristic of acquiescence bias is its temporal stability. Early researchers believed that response styles in general, including ARS, were temporally unstable and should thus be treated as ephemeral responding behavior that might not be generalizable from one survey to another (Hui & Triandis, 1985; Rorer, 1965). However, recent researchers using modern modeling techniques have questioned this

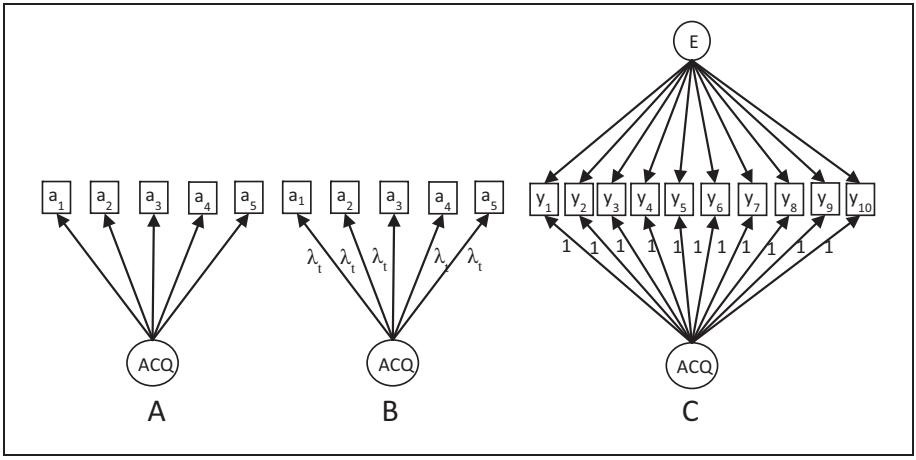

Figure 1. Congeneric factor model (A), (essential) tau-equivalence factor model (B), and measurement model with items that are simultaneously decomposed into a substantive construct factor and a tau-equivalence acquiescence factor (C).

Note. ACQ = acquiescence latent factor; a_1 to a_5 = observed indicators with no measurement of substantive content; λ_t = factor loadings of the acquiescence factor constrained to be identical; y_1 to y_{10} = items measuring substantive construct.

assumption of temporal instability for ARS, by providing direct evidence that ARS is stable both within a survey (Billiet & McClendon, 2000; Weijters et al., 2010a) and across time (Billiet & Davidov, 2008; Weijters et al., 2010b). For instance, by splitting a survey into five successive sets of items, Weijters et al. (2010a) discovered that ARS was consistently and equally present in each set. In particular, all the observed indicators of ARS had identical loadings on the same latent factor, demonstrating that acquiescence bias was consistent within the survey. It has also been observed that acquiescence bias is moderately correlated over spans of 1 year (Weijters et al., 2010b) and 4 years (Billiet & Davidov, 2008). Weijters et al. (2010a) thus concluded that the effect of acquiescence bias could not be dismissed.

The Tau-Equivalence Assumption in Acquiescence Modeling

Operationally, the stability of acquiescence bias can be modeled by a *congeneric* latent factor structure, in which measurement items will load on a single factor of acquiescence within a single administration of a survey (Figure 1A; Weijters et al., 2010a).[1] The factor loadings of this acquiescence latent factor are typically allowed to differ, implying that each item is affected by the latent factor to a varying degree. The observed indicators are also constructed so that they do not share any homogeneous item content (Weijters et al., 2010a), to ensure that the only variance being extracted by the latent factor is participants' response style (i.e., acquiescence).

Some researchers have taken a step beyond the congeneric nature of ARS to propose, as a start that the effect of an acquiescence factor is the same on all measurement items; that is, the factor loadings of the latent acquiescence factor are constrained to be identical for all the measurement items. This is referred to as the (essential) *tau-equivalence assumption*[2] (see Figure 1B; Bollen, 1989; Graham, 2006). In this tau-equivalence model, ARS is assumed to have a consistent and equal effect in overestimating or underestimating the scores of all the measurement items. As shown in Figure 1, similar to Model 1A, Model 1B also requires observed item indicators that maximize content heterogeneity, such that the only common variance between the items is the participants' ARS.

In another variant of the tau-equivalent acquiescence model, all the factor loadings of the acquiescence factor are fixed at 1 (see Figure 1C). The advantage of this model over Model 1B is that it allows simultaneous modeling of (a) a substantive construct factor such as extraversion (for which factor loadings are freely estimated) and (b) an ARS factor with the property of tau-equivalence (for which factor loadings are fixed); this would be impossible with Model 1B for reasons of identification. To avoid the ARS factor extracting any variance of the substantive construct's content, however, the scale in this new model needs to have both regular- and reverse-keyed items (Billiet & McClendon, 2000). The ARS factor would then be able to estimate variances due to the participants' simultaneous agreement (or disagreement) with all the measurement items regardless of keying direction.

Billiet and McClendon (2000) examined the validity of the tau-equivalence assumption of acquiescence (see also Mirowsky & Ross, 1991; Watson, 1992) using 10 survey items relating to political distrust and threat. Each of the 10 measurement items in the model loaded on two latent factors, one represented by a substantive construct and the other by an acquiescence factor. The factor loadings for the substantive construct were freely estimated, whereas those for the acquiescence factor were constrained to be 1 for all the measurement items (i.e., a similar situation to that shown in Figure 1C). In their initial examination, Billiet and McClendon (2000) extracted an acquiescence factor from each of the two scales (political distrust and threat) and discovered that these factors were only moderately correlated ($r = .44$). In a follow-up investigation, they extracted an overall acquiescence latent factor from all 10 items and examined its validity. To obtain an external criterion with which to validate this acquiescence latent factor, Billiet and McClendon (2000) summed the participants' raw scores on the 10 items from the same latent variable model (together with four additional items external to the model). The rationale was that acquiescent participants should agree with most of the items, and thus the ''sum of agreement'' scores should be a reliable reflection of acquiescence bias. The researchers found a strong correlation between the acquiescence latent factor and the external criterion ($r = .90$), supporting the tau-equivalence assumption for acquiescence. However, as Billiet and McClendon (2000) noted, an external criterion is better calculated using items that are independent of the item set already included in the model. Weijters et al. (2010b) also contended that optimal calculation of acquiescence uses measurement items that

are *entirely* external to the model, to ensure that there is no overlap in content between the items in the model and the external criterion. In other words, using (nearly) identical items to model acquiescence and to calculate an external criterion for acquiescence inevitably inflates the correlation between the resulting scores.

Another examination of the tau-equivalence assumption for the acquiescence latent factor was conducted by Weijters et al. (2010a). They randomly chose 112 items from marketing and attitude scales and divided these items into five successive sets of 22 or 23 items. They then calculated an acquiescence score for each set and statistically modeled an overall latent factor of acquiescence using these scores as observed indicators. Weijters et al. (2010a) found that (a) acquiescence had a congeneric nature, with all five acquiescence scores loaded on the same factor; and (b) all the observed acquiescence indicators had the same factor loading on the latent factor, with each indicator having a time-invariant and weak autoregressive relationship with the following indicator across the five successive sets of items. In light of these findings, Weijters et al. (2010a) concluded that the loadings of the different item sets on the common factor were all identical, indicating that ARS was best modeled using a tau-equivalent factor model.

It should be noted, however, that the results of Weijters et al. (2010a) only supported the tau-equivalent assumption for *sets* of items that were heterogeneous in their content. By our interpretation, the findings of Weijters et al. (2010a) were not meant to, and might not, be generalized to individual items. We argue that the tau-equivalent assumption does not necessarily apply at the individual item level. In a typical survey each measurement item has different characteristics, such as item extremity, keying direction, readability, and content ambiguity (McClendon, 1991). Assuming that the study of Weijters et al. (2010a) involved an adequate sample of diverse item types and thoroughly scrambled them, each of these acquiescence item *sets* should have similar characteristics. However, when a study involves individual items only, ARS may interact with the particular characteristics of each item, leading to a disparate effect of response styles on these items. As the validity of the tau-equivalence assumption at the *individual item* level has not been thoroughly examined, the current study was conducted to examine the tenability of this assumption for individual items.

The Present Research

Some doubt has been cast on the efficacy of the tau-equivalence assumption in the modeling of acquiescence bias, despite this type of model being popular in applied research (e.g., Camacho, de Jong, & Stremersch, 2014; Gaylord-Harden, Gipson, Mance, & Grant, 2008). As Geiser, Eid, and Nussbeck (2008) stated, ''All items (positively as well as negatively worded) have equal loadings. . . . This assumption is very strong, as it suggests that there is a homogeneous response style factor that applies in the same way to all items'' (p. 50). As there are insufficient empirical studies examining whether the tau-equivalence assumption is as effective when applied

to item sets as to individual items, the current research examined the validity of this assumption for the acquiescence latent factor at the individual item level. This investigation will advance our understanding of acquiescence bias and, possibly, subsequent statistical modeling of this response style.

To achieve our goal, we first measured an external acquiescence criterion and examined the regression weight of this acquiescence criterion for Big Five personality items and optimism items (Figure 2A). The Big Five personality items were chosen because of the wide use of this scale in many areas of psychology and social science, and optimism items were chosen because recent research has revealed contamination by ARS of the item scores for this measure (e.g., Maydeu-Olivares & Coffman, 2006). If the effect of acquiescence on these items was identical within a measure (i.e., a correct assumption of tau-equivalence), then the regression weights on these measurement items should be identical.

To strengthen our validity examination, we modeled an acquiescence factor on the basis of the tau-equivalence assumption within each substantive measure (i.e., Big Five personality and optimism measures) using SEM. A strong convergent validity would be found (a) if the tau-equivalence acquiescence factor was highly correlated with the external acquiescence criterion (Figure 2B) and (b) if the tau-equivalence acquiescence factors from each measure were strongly correlated (Figure 2C). Finally, we obtained a second-order factor from all the modeled tau-equivalence acquiescence factors and correlated this second-order factor with the external acquiescence criterion (Figure 2D). This method may be used to evaluate the overall evidence for validity of the tau-equivalence assumption.

Several aspects of the current study represented important improvements over previous studies of the tau-equivalence assumption at the item level (e.g., Billiet & McClendon, 2000). Billiet and McClendon (2000) used only two measures to examine their research question, whereas our examination significantly increased the number of measures: five personality scales and one optimism scale. To validate the tenability of the tau-equivalence assumption, we used items that were independent of the SEM model to calculate an external acquiescence score to act as a criterion of validity. In contrast, the tau-equivalence acquiescence factor used as an external criterion of validity by Billiet and McClendon (2000) heavily relied on the same items in their SEM model (i.e., 10 out of the total of 14 acquiescence items are from the model). Finally, to further investigate the validity of our conclusion, we also examined the intercorrelations between various acquiescence factors modeled on the basis of the same tau-equivalence assumption.

Method

Participants and Procedure

One thousand and twenty-three students (664 females, 351 males, and 8 unidentified) from an introductory psychology class at a large Canadian public university completed an online survey in exchange for a course credit. The students had an average

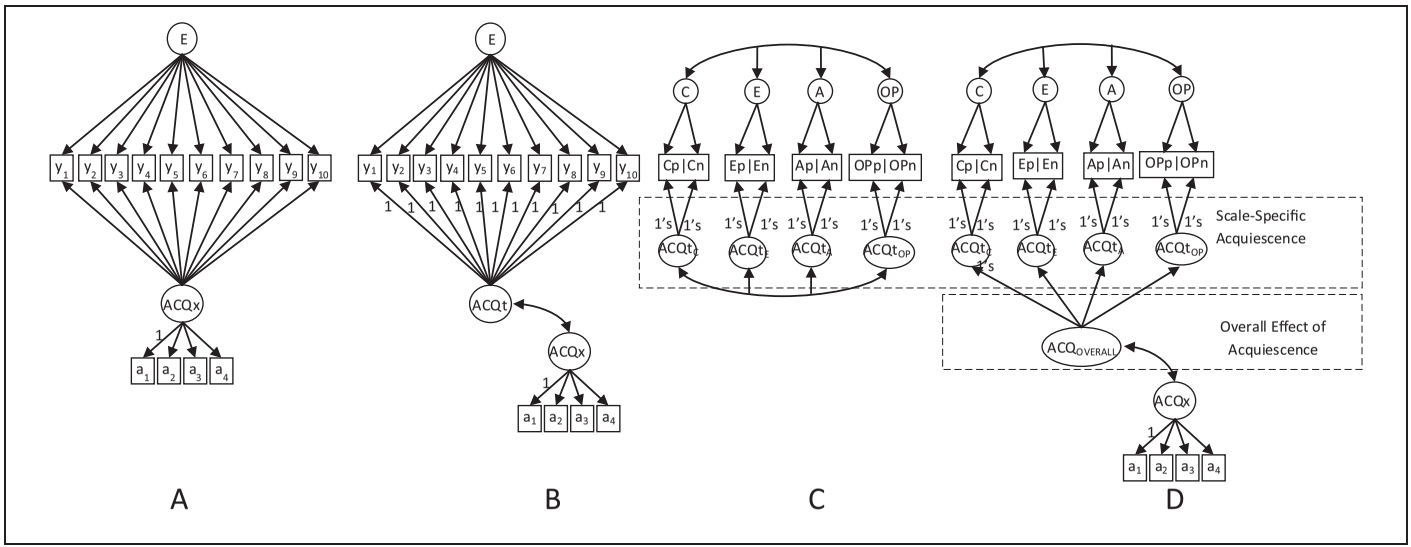

Figure 2. Structural equation models used in this article.

Note. ACQx = acquiescence factor measured by four external indicators (a_1, a_2, a_3, a_4); ACQt = tau-equivalence acquiescence factor extracted from items that simultaneously measure substantive construct factors; y_1 to y_{10} = items measuring substantive constructs; a = acquiescence indicators; OP = optimism; C = conscientiousness; E = extraversion; A = agreeableness; p = a set of regular-keyed items; n = a set of reverse-keyed items.

age of 18.43 years (SD = 1.89). They completed the survey at the beginning of the academic year, and it is thus highly unlikely that these students would have learned about the measures (described below) before taking part in the study. Some of the data were used to examine the nomological network of optimism (Kam & Meyer, 2012), but this work is unrelated to the purpose of the current study.

Measures

International Personality Item Pool (IPIP) Personality Measure. The measure comprised five scales measuring the Big Five personality factors (Goldberg et al., 2006): openness to experience (Cronbach's α = .72, 95% confidence interval [CI] = [.70, .76]), conscientiousness (Cronbach's α = .79, 95% CI = [.76, .81]), extraversion (Cronbach's α = .87, 95% CI = [.85, .90]), agreeableness (Cronbach's α = .76, 95% CI = [.72, .79]), and emotional stability (Cronbach's α = .83, 95% CI = [.80, .85]). Each personality factor was measured using five regular-keyed items and five reverse-keyed items.

Life Orientation Test–Revised (LOT-R). The LOT-R (Scheier, Carver, & Bridges, 1994) has been widely used to measure dispositional optimism. This measure contains six items, half of which are regular-keyed and half reverse-keyed (Cronbach's α = .80, 95% CI = [.77, .83]). Several studies have documented the effects of ARS on this scale (Maydeu-Olivares & Coffman, 2006).

Acquiescence Indicators. We obtained 72 items with diverse content from a wide variety of psychological measures, such as math anxiety, motivation to control prejudice, ''just world'' beliefs, opinions about obese individuals, need for cognition, implicit theory of the world, and other individual difference measures.[3] The interitem correlations were extremely low (r = .03). De Beuckelaer, Weijters, and Rutten (2010; see also Baumgartner & Steenkamp, 2001) suggested that low interitem correlation was a necessary criterion for an acquiescence score because the response style scores of participants should not be contaminated by the measurement of any substantive constructs. Previous research has confirmed the validity of using the averaged item scores from external items as a measurement of acquiescence (Weijters et al., 2010a). Therefore, in preparation for the modeling of the acquiescence latent factor, we randomly divided these 72 items into four sets of 18 items, and averaged the item scores to form one indicator of acquiescence for each set. We later used SEM to model an acquiescence latent factor from the four averaged indicators.

Analysis Strategies

We first tested the assumption that the measurement items had identical factor loadings on the acquiescence latent factor. We allowed the measurement items to load on their respective substantive factors and then externally modeled an acquiescence factor using four external indicators of acquiescence (see Figure 2A). In the first model,

the acquiescence factor was allowed to freely predict each measurement item (e.g., extraversion). In the second model, we then constrained the regression weights of the acquiescence to each measurement item to be identical. If the fit of the second model was worse than that of the first model, the tau-equivalence assumption was not supported. On the contrary, if the fit of the second model was not worse than that of the first model, the tau-equivalence assumption was supported.

In addition to the previous method, we also employed other means to examine our research question. In one method, each measurement item was loaded simultaneously on its substantive factor and on the acquiescence factor. Although the factor loadings for the substantive construct factor were freely estimated, those for the acquiescence factor were fixed at 1 (i.e., assuming tau-equivalence). We then correlated this tau-equivalent acquiescence factor with the acquiescence factor measured by our external indicators (see Figure 2B). A strong correlation would indicate strong convergent validity, supporting the use of the tau-equivalent assumption to approximate acquiescence.

The methods described above used the external acquiescence criterion (i.e., 72 acquiescence items) to validate the tau-equivalence assumption. To strengthen our validity examination, we then examined the intercorrelations between the tau-equivalence acquiescence factors estimated from multiple constructs (Figure 2C). Previous research has found that the degree of acquiescence should be stable within a survey (Weijters et al., 2010a). If the tau-equivalence model reasonably approximated acquiescence, we expected strong correlations between the tau-equivalence acquiescence factors from disparate measures. In the final validity examination for the overall ability of the tau-equivalence acquiescence factor to approximate acquiescence, we constructed a second-order factor model with an overall acquiescence factor above four tau-equivalent first-order acquiescence factors (see Figure 2D). Correlations between the overall acquiescence factor and the external acquiescence criterion were calculated. A strong correlation here would lend support to approximating acquiescence using the tau-equivalence assumption.

All the analyses were conducted using the R system for statistical computing 3.1.0 (R Development Core Team, 2012), with the add-on package lavaan developed by Rosseel (2012) for SEM analysis. Each analysis was conducted using a robust maximum likelihood estimator, which allowed a certain deviation of multivariate normality in the data. Model comparisons were conducted using the appropriate scaled chi-square tests to compare nested models (Satorra, 2000).

Results

Testing Identical Regression Weights for Acquiescence to Individual Items

To test whether the factor loadings for acquiescence are identical between individual measurement items (see Figure 2A), we compared the models with and without assuming tau-equivalence for acquiescence, for each substantive construct (see Tables 1A and 2). Visual examination of the results revealed large variations in the

Table 1. Model Fit.

			χ^2	*df*	*p*	RMSEA	SRMR
(A) Models testing regression weights for tau-equivalence							
Openness	•	Before imposing tau-equivalence	906.42	67	<.001	.11	.08
	•	After imposing tau-equivalence	954.85	76	<.001	.11	.09
Conscientiousness	•	Before imposing tau-equivalence	535.42	67	<.001	.08	.05
	•	After imposing tau-equivalence	556.95	76	<.001	.08	.06
Extraversion	•	Before imposing tau-equivalence	346.28	67	<.001	.06	.04
	•	After imposing tau-equivalence	415.72	76	<.001	.07	.08
Agreeableness	•	Before imposing tau-equivalence	298.90	67	<.001	.06	.05
	•	After imposing tau-equivalence	341.78	76	<.001	.06	.06
Emotional stability	•	Before imposing tau-equivalence	558.62	67	<.001	.09	.05
	•	After imposing tau-equivalence	nc				
Optimism	•	Before imposing tau-equivalence	129.01	29	<.001	.06	.03
	•	After imposing tau-equivalence	179.30	34	<.001	.07	.07
(B) Models testing tau-equivalence acquiescence factors							
Openness			nc				
Conscientiousness			407.35	74	<.001	.07	.05
Extraversion			270.02	74	<.001	.05	.04
Agreeableness			209.36	74	<.001	.04	.04
Emotional stability			596.33	74	<.001	.08	.06
Optimism			80.536	32	<.001	.04	.03
(C) Models allowing tau-equivalence acquiescence factors to be intercorrelated			1834.51	578	<.001	.05	.07
(D) Second-order factor model testing overall ability of tau-equivalence acquiescence factors to approximate acquiescence			2109.21	721	<.001	.04	.07

Note. N = 1,023. nc = nonconvergence.

standardized regression weight for the items in each examined scale. Constraining the unstandardized regression weights (from the acquiescence factor to all items) to be identical caused the fit of the models to be significantly worse in all scales (all p values were $<.001$, according to scaled chi-square difference tests; Satorra, 2000), with the constrained model for emotional stability even failing to reach convergence. The effect of acquiescence on the individual items therefore differed significantly, demonstrating that respondents might not uniformly or indiscriminately acquiesce on all items.

Validity of the Tau-Equivalence Acquiescence Factor

Although the effects of acquiescence on individual items were not identical, we were nevertheless interested in the ability of the acquiescence factor modeled on the basis of assumed tau-equivalence to approximate the response style. If the ability was strong, the tau-equivalence assumption would probably still be reasonable even if it might not hold strictly. We therefore correlated the tau-equivalence acquiescence factor with the external acquiescence criterion (see Figure 2B). The results are shown in Tables 1B and 3. In general, the external acquiescence criterion was more strongly correlated with the tau-equivalence acquiescence factors (rs = .33-.80) than with the substantive construct factors (rs = .15-.28), suggesting some ability of the tau-equivalence acquiescence factors to capture the response style. Although the strongest convergent validity evidence was found for emotional stability (r = .80), that for most of the other constructs was much weaker (r = .33-.52). Therefore, for most of the constructs in the current investigation, the tau-equivalence acquiescence factor only demonstrated a weak to moderate validity.

We next examined the correlations between the tau-equivalence acquiescence factors (see Figure 2C). We simultaneously modeled six construct factors and six corresponding tau-equivalence acquiescence factors from the Big Five personality and optimism items. The results failed to achieve satisfactory convergence. Further investigation revealed nonconvergence in the ''openness to experience'' and ''emotional stability'' scales. We therefore excluded these two scales from subsequent analyses. The results in Table 4 demonstrate some evidence for convergent validity, with correlations for the acquiescence factors between the scales ranging from .32 to .73. The correlation (r = .73) between the acquiescence factors of the ''agreeableness'' and ''conscientiousness'' scales was the strongest, and our previous analysis had also demonstrated reasonably good convergent validity for the acquiescence factors of these two constructs with external acquiescence criterion (r = .52 for ''agreeableness'' and .47 for ''conscientiousness''; see Table 3). In other words, acquiescence factors that demonstrated stronger evidence of validity in our previous analysis also demonstrated higher inter-measure convergent validity here.

Finally, we modeled an overall acquiescence factor atop of the four tau-equivalence first-order acquiescence factors from the previous analysis (see Figure 2D). The overall acquiescence factor was moderately correlated with the external

Table 2. Unstandardized (Standardized) Regression Weights of Acquiescence Factors Before the Assumption of Tau-Equivalence.

	O	C	E	A	ES	Optimism[a]
Regular 1	1.89* (.08)	5.32*** (.25)	6.15*** (.33)	4.71*** (.25)	−2.75*** (−.12)	1.37*** (.22)
Regular 2	3.38** (.16)	3.36*** (.19)	6.33*** (.32)	4.75*** (.26)	−3.26*** (−.15)	1.36*** (.26)
Regular 3	0.66 (.03)	1.19*** (.18)	6.31*** (.32)	4.46*** (.30)	−0.95 (−.04)	0.96 (.20)
Regular 4	4.35*** (.24)	4.99*** (.31)	4.66*** (.21)	4.17*** (.24)	1.34 (.06)	—
Regular 5	4.15*** (.27)	5.14*** (.28)	5.63*** (.30)	5.77*** (.36)	2.69** (.11)	—
Reverse 1	0.38 (.02)	0.49 (.02)	−2.01* (−.09)	3.43 (.16)	1.00 (.04)	0.05 (.01)
Reverse 2	−1.23 (−.05)	1.07 (.05)	−2.78*** (−.13)	−0.16 (−.01)	2.86*** (.13)	−0.42* (−.08)
Reverse 3	−0.46 (−.02)	2.20* (.09)	−2.64*** (−.12)	1.31 (.07)	4.48*** (.23)	−0.62** (−.12)
Reverse 4	1.00 (.04)	1.16 (.06)	−1.53* (−.07)	1.72* (.08)	−0.37 (−.02)	—
Reverse 5	0.59 (.03)	1.46* (.08)	−1.89* (−.08)	−0.38 (−.02)	4.50*** (.23)	—

Note. N = 1,023. Regular = regular-keyed items; Reverse = reverse-keyed items; O = openness to experience; C = conscientiousness; E = extraversion; A = agreeableness; ES = emotional stability.

a. The optimism scale contains four regular-keyed items and three reverse-keyed items only.

*$p < .05$. **$p < .01$. ***$p < .001$.

Table 3. Correlation of Externally Formed Acquiescence With Substantive Latent Constructs and Tau-Equivalence Acquiescence Latent Factors.

	Correlation with substantive construct	Correlation with tau-equivalence acquiescence factor
Openness	nc	nc
Conscientiousness	.15***	.47***
Extraversion	.28***	.41***
Agreeableness	.19***	.52***
Emotional stability	.19***	.80***
Optimism	.20***	.33***

Note. N = 1,023. nc = nonconvergence of the model.
***$p < .001$.

acquiescence criterion, $r = .58$, $z = 7.03$, $p < .001$, demonstrating satisfactory but not strong validity. As the validity correlations of the tau-equivalence (first-order) acquiescence factors only had values between .33 (for optimism) and .52 (for agreeableness) when the four individual measures were assessed separately (see Table 3), the overall acquiescence factor derived simultaneously from these four measures showed slightly higher validity (.58).

Investigating the Strong Validity Coefficient for Emotional Stability

When the externally measured acquiescence criterion was correlated with the tau-equivalence acquiescence factor (Figure 2B), we found the validity coefficient to be substantially stronger for emotional stability ($r = .80$) than for comparative measures ($r = .33$-$.52$; Table 3). We investigated this issue further to try to shed light on why the validity of the tau-equivalence assumption varied from measure to measure. A closer examination of the emotional stability scale showed that many of its items involved parallel content. The item ''I seldom feel blue'' had nearly identical wording to its reverse-keyed counterpart ''I often feel blue.'' The item ''I dislike myself'' had similar wording to items with antithetical meanings: ''I feel comfortable with myself'' and ''I am very pleased with myself.'' Obviously, participants who simultaneously endorsed antithetical items (e.g., ''I seldom feel blue'' and ''I often feel blue'') were likely to be demonstrating ARS. Indeed, researchers have often summed participants' scores from such antithetical items to use as a proxy for ARS (e.g., Rammstedt & Farmer, 2013). Given the large number of items of this nature contained in this measure, the emotional stability scale was probably affected by ARS more than other scales in the current study. Logically, if these antithetical item pairs can effectively capture acquiescence, then adding residual covariances within each item pair would substantially attenuate their effect on the acquiescence factor. Indeed, the validity of the tau-equivalence acquiescence factor with the external acquiescence criterion dropped dramatically, from .80 to .47, after we did so. The

Table 4. Correlations Between Construct Factors and Between Tau-Equivalence Acquiescence Factors.

	1.	2.	3.
Construct factors			
1. Conscientiousness			
2. Extraversion	.29***		
3. Agreeableness	.35***	.23***	
4. Optimism	.43***	.44***	.28***
Acquiescence factors from			
1. Conscientiousness			
2. Extraversion	.50***		
3. Agreeableness	.73***	.66***	
4. Optimism	.35***	.30***	.32***

Note. N = 1,023.
***$p < .001$.

residual covariances greatly reduced the amount of acquiescence captured by the tau-equivalence factor, confirming our intuition that when the developer of a scale includes within it (nearly) identical items with opposite wording, these item pairs should be particularly effective in capturing acquiescence.[4]

Discussion and Research Recommendations

The purpose of the current study was to examine the tau-equivalence assumption in the modeling of acquiescence. We conducted this examination by (a) correlating the tau-equivalence acquiescence factor with acquiescence measured by external indicators and (b) correlating the tau-equivalence acquiescence factors (from various scales) with one another. We found that acquiescence measured by external indicators was more strongly correlated with the tau-equivalence acquiescence factors (r = .33-.80) than with substantive constructs such as extraversion and conscientiousness (r = .15-.28). For the tau-equivalence acquiescence factor of the emotional stability scale, the correlation was as high as .80, providing evidence of convergent validity. We also found good convergent validity for the overall tau-equivalence acquiescence factor (r = .58). These results support, to some degree, the validity of measuring acquiescence on the basis of the tau-equivalence assumption.

Some of the evidence found was less supportive of the validity of the tau-equivalence assumption. First, the regression weights for the acquiescence factor (measured by external indicators) were not strictly identical for different scale items. The validity of the assumption also varied widely between the scales, indicating that violation of the tau-equivalence assumption might vary from one scale to another. All the validity evidence was weaker than that found in Billiet and McClendon's (2000) study. Unlike Billiet and McClendon (2000), we did not use raw items

extracted from the measurement model to calculate the external criterion for acquiescence, but instead averaged 72 external items (that did not overlap with the items in our SEM model) for this purpose. We therefore believe that our results provide a more accurate picture of the validity of the tau-equivalence assumption in the measurement of acquiescence.

From a methodological perspective, when interpreted with the findings of Weijters et al. (2010a), the current study clarifies the situation in which the tau-equivalence assumption of acquiescence holds. As Weijters et al. (2010a) were interested in the general stability of acquiescence throughout an entire survey, they formed parcels (aggregates) of acquiescence indicators based on 22 or 23 items from a diverse array of scales in different content areas. Their strong evidence for tau-equivalence among these indicators implies that the participants exerted a similar degree of acquiescence bias at the *aggregated* level throughout the survey. Our study, in contrast, examined the degree of acquiescence bias at the *individual item* level. Our results indicate that the tau-equivalence assumption does not hold at the individual item level, implying that acquiescence interacts differently with the characteristics of each individual item (De Beuckelaer et al., 2010). It should therefore not be assumed that tau-equivalence proven at the aggregated level necessarily applies at the individual item level.

Regarding a possible interaction between acquiescence and individual item characteristics, the current study revealed a situation in which such an interaction might occur. The emotional stability scale used in the current study included several items that had similar wording but opposite keying directions. When we allowed the antithetical items within each of these pairs to covary with each other (which would substantially reduce their effects on the acquiescence factor), the validity of the tau-equivalence acquiescence factor was dramatically reduced. This result was consistent with our intuition that the sum-of-agreement scores from items of antithetical meaning would capture ARS (e.g., Rammstedt & Farmer, 2013). If the developer of a scale mixes antithetical items with other items, the tau-equivalence assumption will naturally be violated because the antithetical item pairs will be more effective in capturing ARS than the other items. This finding alone can explain why the tau-equivalence assumption does not necessarily hold for individual items.

Overall, our results suggest that modeling ARS on the basis of the tau-equivalence assumption has a certain degree of validity although this assumption does not hold at the individual item level. In light of this finding, if a situation arises in which a researcher cannot use external items to calculate acquiescence, using the tau-equivalence assumption to model the acquiescence latent factor may be a reasonable alternative. Our results suggest that there is some support for the validity of this assumption overall, although the specific validity of the method varied significantly from one scale to another (as shown in Tables 3 and 4).

When researchers have the option of using an explicit method to directly measure acquiescence, we believe that such explicit methods are preferable. A method recommended by De Beuckelaer et al. (2010) uses items with heterogeneous content and

averages participants' agreement with these items. Careful planning is required before using this method, as it requires a set of dedicated items that measure diverse item content. Interitem correlations should also be low (De Beuckelaer et al., 2010), to ensure that the ''sum of agreement'' scores will not measure any substantive constructs. To model acquiescence using the SEM approach, Weijters et al. (2010) used sets of item aggregates from heterogeneous items as multiple observed indicators, all loaded on an acquiescence latent factor. This has the practical advantage of avoiding a formidable number of observed indicators in the model (De Beuckelaer et al., 2010a), and we followed this approach when modeling the external acquiescence factor in the current study (see Figure 2A). In the current study, we used 72 items to measure this external acquiescence criterion. From the psychometric perspective, a larger number of items may provide a more stable measurement of the acquiescence construct (Weijters et al., 2010b).

At least two other methods may be used to control for ARS. One such method relies on the use of an advanced item response theory model. This method probably requires a larger sample size, but can theoretically control not only for acquiescence but also for other response styles such as extreme responding and moderate responding (Bolt et al., 2014). Another method of controlling for acquiescence is to include a subset of items in an SEM model that are unlikely to be contaminated by the response style. For example, Billiet, Cambré, and Welkenhuysen-Gybels (2002) included an item that was measured in a dichotomous format (i.e., ''Do you think the number of immigrants to [country] nowadays should be . . . [choose an answer of 'increased' or 'reduced']'') among other items measured using a Likert-type scale. As this item was probably not affected by acquiescence, it loaded on the same content factor as the other items but not on the acquiescence factor. This SEM model is identified even without the tau-equivalence assumption for the acquiescence factor. Note that all these methods require careful research design before data collection can take place.

A final question relates to the status or nature of the tau-equivalence response style factor: What is it measuring? This response style factor appears to measure not only acquiescence but also other constructs. One possible candidate is respondents' differential thresholds for each scale anchor (Maydeu-Olivares & Coffman, 2006). A scale anchor such as ''moderately agree'' is usually ambiguous in its meaning. One respondent may interpret the anchor as slight agreement, whereas another may interpret it as a stronger agreement. Based on the opinion by Maydeu-Olivares and Coffman (2006), the tau-equivalence response style factor may partially control for the subjective interpretation of scale anchors by respondents, although its exact nature deserves to be further explored.

The current study has some limitations. First, although our results suggest that acquiescence may interact with item characteristics, how this occurs has not been investigated. Although demonstrating such an interaction is of itself important, future research should explore why the interaction may occur. For example, some researchers have suggested that participants may be more likely to demonstrate acquiescence

if the item content is ambiguous or difficult (Hanley, 1965; but see McClendon, 1991, for negative results). Second, the form of acquiescence measured in this study was acceptance acquiescence, which, according to Bentler, Jackson, and Messick (1971), is distinct from agreement acquiescence. If acceptance acquiescence was the only determinant of item responses, the pattern of correlations between the items would be different from those of agreement acquiescence (Bentler et al., 1971; Table 2). Hence, different types of ARS should be considered in future investigations. Third, although the current study examined the effect of acquiescence at the individual item level and proposed a possible acquiescence × item interaction, previous studies have suggested that the ARS may also interact with person-level characteristics. For example, Knowles and Condon (1999) found that acquiescent respondents tended to truncate their reconsideration of an item content before responding. Taken with the present study, this previous finding suggests that the possible acquiescence × item interaction may be moderated by person-level characteristics, and future research should continue to clarify the intricate interactions between these factors. Moreover, the participants in the current study were university students, who could be different from general populations in terms of response styles. We also only examined the results with a Canadian sample. Future research may use samples from different countries (e.g., Welkenhuysen-Gybels, Billiet, & Cambré, 2003) to enhance the generalizability of the results. Finally, the tau-equivalence acquiescence model of the current study failed to achieve convergence for some construct measures (e.g., openness). In our experience, failure to achieve convergence with this type of model is not uncommon. The current research opens the door to investigating convergence issues with this type of model.

Appendix

Explanations of the Congeneric and Tau-Equivalence Models in Matrix Form

Modeling of Congeneric Model

Assume that no substantive factors other than the acquiescence factor are measured by these items. Algebraically, the response of participant j for item i (denoted by y_{ij}) can be written as

$$y_{ij} = \mu_i + \lambda_i \zeta_j + e_{ij}, \quad (1)$$

where μ_i is the intercept specific for item i, λ_i is the factor loading for item i, ζ_j is the acquiescence factor score for participant j that is common across all of the items, and e_{ij} is the residual for participant j for item i. The equation is very similar to that of the common one-factor model, with the factor score ζ here representing a response style rather than a substantive construct. Note that in a congeneric model, the factor loading λ is identical for an item *across all participants*, but each item has a different λ value. Assume that there is a total of p items in the model. In compact matrix form, the equation becomes

$$\mathbf{y} = \boldsymbol{\mu} + \boldsymbol{\Lambda}_a \zeta + \mathbf{e}, \tag{2}$$

where $\mathbf{y}$, $\boldsymbol{\mu}$, $\boldsymbol{\Lambda}_a$, and $\mathbf{e}$ are $p \times 1$ vectors and ζ is a scalar representing the acquiescence factor score (see Figure 1A).

Modeling of Tau-Equivalence Models

A tau-equivalence model goes beyond a congeneric model by assuming that all the factor loadings in Equation 2 (all of the λ values or all of the elements in $\boldsymbol{\Lambda}_a$) have equal values (see Figure 1B).

In applied research, investigators are often mainly interested in substantive constructs and the acquiescence factor is only included to partial out the effect of the ARS (e.g., Aichholzer, 2014). However, the simultaneous modeling of a substantive construct and the acquiescence factor is not possible in the absence of an explicit measurement of acquiescence, and a method for modeling acquiescence without explicitly measuring this response style has therefore been developed. Typically, one substantive construct is included in the model together with one acquiescence factor (Figure 1C):

$$\mathbf{y} = \boldsymbol{\mu} + \boldsymbol{\Lambda}_b \eta + \mathbf{1}\zeta + \mathbf{e}, \tag{3}$$

where $\mathbf{y}$, $\boldsymbol{\mu}$, $\boldsymbol{\Lambda}_b$, $\mathbf{e}$ and $\mathbf{1}$ are $p \times 1$ vectors. η refers to the substantive construct, with its factor loading matrix $\boldsymbol{\Lambda}_b$. Both η and ζ are scalars. The substantive factor η is expected to have a mean of zero and a variance of 1. All the factor loadings for the acquiescence factor ζ have the value 1 and ζ is expected to have a mean of zero and a freely estimated variance. The acquiescence factor ζ is often assumed to be uncorrelated with the substantive factor η. Equation 3 can also be expressed as follows:

$$\mathbf{y} = \boldsymbol{\mu} + \boldsymbol{\Lambda}^* \boldsymbol{\eta}^* + \mathbf{e}, \tag{4}$$

where $\boldsymbol{\Lambda}^* = [\boldsymbol{\Lambda}_b \mid \mathbf{1}]$ and $\boldsymbol{\eta}^* = [\eta \ \zeta]'$. $\boldsymbol{\Lambda}^*$ is a $p \times 2$ matrix of factor loadings and $\boldsymbol{\eta}^*$ is a 2×1 vector of latent factors, namely, the substantive factor η and the acquiescence factor ζ; these factors are assumed to be uncorrelated for identification purposes. This model can be easily assessed using SEM programs. The model can also be extended to include multiple substantive factors rather than just one.

Acknowledgment

We would like to thank Xitao Fan for his feedback to an earlier version of this article.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported financially by Start-up Research Grant by the University of Macau awarded to Chester Chun Seng Kam.

Notes

1. See the appendix for a discussion of the assumption in matrix form.
2. Unlike strict tau-equivalence, essential tau-equivalence allows the item intercept to differ between all the measurement items when the items are loaded on the acquiescence latent factor (Graham, 2006; Jackman, 1973). Following the usual practice in the literature (e.g., Traub, 1994), in this article we simply refer to essential tau-equivalence as tau-equivalence.
3. The items used for the measurement of acquiescence are available from the first author.
4. In addition to the tau-equivalence assumption, Weijters et al. (2010a) found that each item set had an autoregressive relationship with the following item set within the same survey. We attempted to model this autoregressive relationship, with an item predicting the one following it within a scale. This analysis was possible because each item within a personality subscale appeared consistently once every five questions and because the optimism items were adjacent to each other. The autoregressive models either did not converge properly or gave noninterpretable results. We therefore believe that this autoregressive relationship may only be relevant with item aggregates (as in Weijters et al., 2010a) and is untenable for individual items.

References

Aichholzer, J. (2014). Random intercept EFA of personality scales. *Journal of Research in Personality*, *53*, 1-4.

Baumgartner, H., & Steenkamp, J. B. E. (2001). Response styles in marketing research: A cross-national investigation. *Journal of Marketing Research*, *38*, 143-156.

Bentler, P. M., Jackson, D. N., & Messick, S. (1971). Identification of content and style: A two-dimensional interpretation of acquiescence. *Psychological Bulletin*, *76*, 186-204.

Billiet, J. B., & Davidov, E. (2008). Testing the stability of an acquiescence style factor behind two interrelated substantive variables in a panel design. *Sociological Methods & Research*, *36*, 542-562.

Billiet, J. B., & McClendon, M. J. (2000). Modeling acquiescence in measurement models for two balanced sets of items. *Structural Equation Modeling*, *7*, 608-628.

Billiet, J., Cambré, B., & Welkenhuysen-Gybels, J. (2002). Equivalence of measurement instruments for attitude variables in comparative surveys taking method effects into account: The case of ethnocentrism. In A. Ferligoj & A. Mrvar (Eds.), *Developments in social science methodology* (pp. 73-96). Ljubljana, Slovenia: FDV.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: Wiley.

Bolt, D. M., Lu, Y., & Kim, J.-S. (2014). Measurement and control of response styles using anchoring vignettes: A model-based approach. *Psychological Methods*. Advance online publication. doi:10.1037/met0000016

Camacho, N., de Jong, M., & Stremersch, S. (2014). The effect of customer empowerment on adherence to expert advice. *International Journal of Research in Marketing*, *31*, 293-308.

Darcy, M., & Tracey, T. J. (2003). Integrating abilities and interests in career choice: Maximal versus typical assessment. *Journal of Career Assessment*, *11*, 219-237.

De Beuckelaer, A., Weijters, B., & Rutten, A. (2010). Using ad hoc measures for response styles: A cautionary note. *Quality & Quantity*, *44*, 761-775.

Friborg, O., Martinussen, M., & Rosenvinge, J. H. (2006). Likert-based vs. semantic differential-based scorings of positive psychological constructs: A psychometric comparison of two versions of a scale measuring resilience. *Personality and Individual Differences*, *40*, 873-884.

Gaylord-Harden, N. K., Gipson, P., Mance, G., & Grant, K. E. (2008). Coping pattern of African American adolescents: A confirmatory factor analysis and cluster analysis of the children's coping strategies checklist. *Psychological Assessment*, *20*, 10-22.

Geiser, C., Eid, M., & Nussbeck, F. W. (2008). On the meaning of the latent variables in the CT-C (M-1) model: A comment on Maydeu-Olivares and Coffman (2006). *Psychological Methods*, *13*, 49-57.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, *40*, 84-96.

Graham, J. M. (2006). Congeneric and (essentially) tau-equivalent estimates of score reliability: What they are and how to use them. *Educational and Psychological Measurement*, *66*, 930-944.

Hanley, C. (1965). Personality item difficulty and acquiescence. *Journal of Applied Psychology*, *49*, 205-208.

Hui, C. H., & Triandis, H. C. (1985). The instability of response sets. *Public Opinion Quarterly*, *49*, 253-260.

Jackman, M. R. (1973). Education and prejudice or education and response-set?*American Sociological Review*, *38*, 327-339.

Jackson, D. N., & Messick, S. (1962). Response styles on the MMPI: Comparison of clinical and normal samples. *Journal of Abnormal and Social Psychology*, *65*, 285-299.

Kam, C., & Meyer, J. P. (2012). Do optimism and pessimism have different relationships with personality dimensions? A re-examination. *Personality and Individual Differences*, *52*, 123-127.

Knowles, E. S., & Condon, C. A. (1999). Why people say ''yes'': A dual-process theory of acquiescence. *Journal of Personality and Social Psychology*, *77*, 379-386.

Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors?*Journal of Personality and Social Psychology*, *70*, 810-819.

Maydeu-Olivares, A., & Coffman, D. L. (2006). Random intercept item factor analysis. *Psychological Methods*, *11*, 344-362.

McClendon, M. J. (1991). Acquiescence: Tests of the cognitive limitations and question ambiguity hypotheses. *Journal of Official Statistics*, *7*, 153-166.

Mirowsky, J., & Ross, C. E. (1991). Eliminating defence and agreement bias from measures of the sense of control: A 22 index. *Social Psychology Quarterly*, *54*, 127-145.

R Development Core Team. (2012). *R: A language and environment for statistical computing* [Computer software manual]. Vienna, Austria: Author.

Rammstedt, B., & Farmer, R. F. (2013). The impact of acquiescence on the evaluation of personality structure. *Psychological Assessment*, *25*, 1137-1145.

Rammstedt, B., Kemper, C. J., & Borg, I. (2013). Correcting Big Five personality measurements for acquiescence: An 18-country cross-cultural study. *European Journal of Personality*, *27*(1), 71-81.

Rorer, L. G. (1965). The great response-style myth. *Psychological Bulletin*, *63*, 123-156.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1-36.

Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In R. D. H. Heijmans, D. S. G. Pollock, & A. Satorra (Eds.), *Innovations in multivariate statistical analysis. A Festschrift for Heinz Neudecker* (pp. 233-247). London, England: Kluwer Academic.

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of the Life Orientation Test. *Journal of Personality and Social Psychology*, *67*, 1063-1078.

Schweizer, K. (2012). A weighted version of the tau-equivalent model of measurement for items with ordered response categories. *International Journal of Statistics and Probability*, *1*, 151-163.

Traub, R. E. (1994). *Reliability for the social sciences: Theory and applications* (Vol. 3). Thousand Oaks, CA: Sage.

Watson, D. (1992). Correcting for acquiescent response bias in the absence of a balanced scale: An application to class consciousness. *Sociological Methods & Research*, *21*, 52-88.

Weijters, B., Geuens, M., & Schillewaert, N. (2010a). The individual consistency of acquiescence and extreme response style in self-report questionnaires. *Applied Psychological Measurement*, *34*, 105-121.

Weijters, B., Geuens, M., & Schillewaert, N. (2010b). The stability of individual response styles. *Psychological Methods*, *15*, 96-110.

Welkenhuysen-Gybels, J., Billiet, J., & Cambré, B. (2003). Adjustment for acquiescence in the assessment of the construct equivalence of Likert-type score items. *Journal of Cross-Cultural Psychology*, *34*, 702-722.

Wetzel, E., Carstensen, C. H., & Böhnke, J. R. (2013). Consistency of extreme response style and non-extreme response style across traits. *Journal of Research in Personality*, *47*, 178-189.

Wiggins, J. S. (1973). *Personality and prediction: Principles of personality assessment*. Reading, MA: Addison-Wesley.