Twin Research and Human Genetics
Volume 16 ▌ Number 4 ▌ pp. 790–798 ▌ © The Authors 2013 ▌ doi:10.1017/thg.2013.38

Heritability of Acquiescence Bias and Item Keying Response Style Associated With the HEXACO Personality Scale

Chester Kam,[1] Julie Aitken Schermer,[2] Juliette Harris,[3] and Philip A. Vernon[4]
[1]*Faculty of Education, The University of Macau, Macau SAR, China*
[2]*Management and Organizational Studies, The University of Western Ontario, London, Ontario, Canada*
[3]*Department of Twin Research and Genetic Epidemiology, King's College London, London, UK*
[4]*Department of Psychology, The University of Western Ontario, London, Ontario, Canada*

The current research investigates the heritability of two of the most common response styles: acquiescence bias (tendency to agree or disagree with survey items regardless of the items' actual content) and item keying (differential responding related to the use of regular- and reverse-keyed items). We estimated response styles from a common personality measure (HEXACO) and examined the heritability of each with univariate genetics analyses. The results show item keying effect was heritable but acquiescence bias was not. Neither response style was strongly influenced by the shared environment of the twins. Unique environmental effects were found to be substantial for response styles. The current findings have important implications for future research of response behaviors that are often overlooked by behavioral geneticists.

■ **Keywords:** personality, twin study, acquiescence, item keying, heritability

People tend to demonstrate substantial differences in how they respond to a typical survey item, and the array of these response styles might suggest certain important biological, physiological, and cultural characteristics of these individuals (Jackson & Messick, 1958). When a person consistently agrees to survey items regardless of their content, for example, their behaviors may demonstrate the manifestation of certain substantive traits. A traditional assumption concerning response bias is that it constitutes transient characteristics caused entirely by situational irregularities (e.g., Rorer, 1965). However, recent research in this area consistently suggests that response styles can be temporally stable (e.g., Alessandri et al., 2010; Couch & Keniston, 1960; Marsh et al., 2010; Motl & DiStefano, 2002; Quilty et al., 2006). Our current research contributes to this area of study by examining the heritability of two of the most important response styles: acquiescence bias and item keying effect. If both acquiescence bias and item keying effect are heritable, then both may represent some substantive biological characteristics that are worthy of investigations on its own. Answers to the nature versus nurture question with regard to response styles will provide opportunities for future research studies on their etiology.

Acquiescence Bias and Item Keying Effect

Cronbach (1942) was one of the first researchers to popularize the issue of acquiescence bias (see also Block, 1965; Lentz, 1938; Lorge, 1937). When administering true–false tests with a group of students, Cronbach observed that some test takers were more likely to choose 'true' as an answer to an item compared to other test takers. Acquiescence bias refers to a participant's tendency to agree or disagree with survey items regardless of the items' actual content. Cronbach's (1941, 1942) seminal discovery on this response style has spawned decades of research in this area, including but not limited to the nature of acquiescence bias (Heal & Sigelman, 1995), its stability within a survey and across time (Billiet & McClendon, 2000; Weijters et al., 2010), its prevalence (Messick & Jackson, 1961), and its cross-cultural manifestations (Grimm & Church, 1999; Kam et al., 2012).

RECEIVED 12 April 2013; ACCEPTED 13 May 2013.

ADDRESS FOR CORRESPONDENCE: Chester Kam, Silver Jubilee Building, The University of Macau, Av. Padre Tomas Pereira, Taipa, Macau SAR, China. E-mail: chesterkam@umac.mo

In one of the early studies on acquiescence bias, Jackson and Messick (1958) discovered that many participants agree with logically contradictory statements (i.e., regular- vs. reverse-keyed items; see also Messick & Jackson, 1961, for similar findings). Later researchers often regarded acquiescence bias as an error caused by cognitive impairment, low verbal ability, poor education, low social status, and cognitive aging (e.g., Ayidiya & McClendon, 1990; Krosnick et al., 1996; Lenski & Leggett, 1960; Meisenberg & Williams, 2008; Messick & Frederiksen, 1958, 1961; Narayan & Krosnick, 1996; Rammstedt et al., 2010; Sigelman et al., 1981). These findings subsequently suggest that acquiescence bias was an error to be controlled for.

To mitigate the effect of acquiescence bias, researchers often include reverse-keyed items within questionnaires (e.g., Rammstedt, et al., 2010). Regular-keyed items measure the presence of a construct whereas reverse-keyed items measure the absence of a construct. When participants agree to both regular-keyed items and reverse-keyed items simultaneously, the acquiescence bias in regular-keyed and reverse-keyed items balances out. Ironically, attempts to control for acquiescence bias often unintentionally engender the second type of response style — item keying effect (Kam & Meyer, 2012; Marsh, 1996; Tomás & Oliver, 1999).

According to Kam and Meyer (2012), regular- and reverse-keyed items usually differ in the valence or the favorability of their content (Peabody, 1967). For instance, for the construct conscientiousness, regular-keyed items (e.g., 'I am careful') are more favorable in content compared to reverse-keyed items (e.g., 'I am careless'). Participants may answer a questionnaire item based on the valence of the item in addition to its content. Item keying effect thus refers to participants' differential responding styles between regular- and reverse-keyed items (Kam & Meyer, 2012). In a recent study, Biderman et al. (2011) found the existence of a small item keying effect even in Big Five personality measures. Given that the Big Five personality measure is widely used, the findings speak about the prevalence of the item keying effect in a great deal of psychological research.

Stability Nature of the Two Response Styles

Research has shown that both acquiescence bias and item keying effect may represent certain stable characteristics. Weijters et al. (2010) recently modeled acquiescence bias with structural equation modeling (SEM). These authors split a survey into five successive sets and discovered that acquiescence bias consistently and equally existed in each set. Particularly, all observed indicators of acquiescence bias had identical loadings on the same latent factor in their SEM investigation, demonstrating that acquiescence bias is consistent within a survey. In another study, Billiet and McClendon (2000) found that the acquiescence latent factor in an SEM model correlates strongly with another operationalization of acquiescence bias ('sum of agreements' on survey items; $r = 0.90$), thus supporting the convergent validity of modeling acquiescence bias with SEM. Based on Weijters et al.'s (2010) research, the current investigation will assume the magnitude of acquiescence bias to be identical among different items within the same survey. Finally, researchers have discovered that acquiescence bias from the same survey correlated at around 0.6 over a span of four years (Bachman & O'Malley, 1984; Billiet & Davidov, 2008) and over four 1-year periods (Marsh et al., 2010; Motl & DiStefano, 2002).

In addition to temporal stability, there is some evidence that item keying response style actually represents a heritable quality. Alessandri et al. (2010) extracted response style factors associated with regular- and reverse-keyed items in an optimism measure and compared these factors between monozygotic (MZ) and dizygotic (DZ) twins. These authors discovered that MZ twins have substantially higher item keying effect than DZ twins. This result indicates that genetics may play a role in causing the higher correlation of an item keying effect among MZ twins. In addition, these authors estimated the heritability of item keying effects to be 20% for optimism items but only 6% for pessimism items. The heritability estimate was statistically significant for optimism items but non-significant for pessimism items. The researchers did not find that shared environmental influences predicted the item keying effect. Alessandri et al. (2010) thus concluded that the item keying effect was found heritable for optimism items that were also positively valenced, but not for pessimism items that were also negatively valenced.

The conclusions by Alessandri et al. (2010), however, can be challenged on several grounds. First, some researchers have argued that optimism and pessimism are two correlated but distinct constructs. Therefore, any item keying effect extracted from optimism items (regular-keyed items) would represent a construct of optimism but not pessimism. Critics may thus interpret Alessandri et al.'s (2010) findings as the construct optimism being more heritable than the construct pessimism, rather than item keying effect being heritable. Second, Alessandri et al. (2010) only examined one psychological measure in their study. Therefore, any interpretation of their findings might be confined only to the construct of optimism and their results may not be generalized to other constructs such as personality traits. Third, Alessandri et al. (2010) studied only item keying response style. Their conclusion may differ substantially when acquiescence bias is also estimated simultaneously with item keying response effect. Our empirical investigations, detailed in the next section, will help build upon this area of research.

The Current Research

The current study aims to advance the research of Alessandri et al. (2010) in several ways. First, past research usually estimated either acquiescence bias (e.g., Weijters et al., 2010) or item keying effect (e.g., Alessandri et al., 2010; Motl & DiStefano, 2002) but seldom assessed them simultaneously

https://doi.org/10.1017/thg.2013.38 Published online by Cambridge University Press

TABLE 1
Model Comparisons

	Model fit						Model comparisons		
	χ^2 (*df*)	*p*	TLI	CFI	RMSEA	SRMR	Comparisons	$\Delta\chi^2$ (Δdf)	*p*
$M_{baseline}$	3878.00 (362)	<.001	0.77	0.80	0.06	0.06			
$M_{positive}$	2664.02 (348)	<.001	0.84	0.87	0.05	0.06	$M_{positive}$ vs. $M_{baseline}$	1076.12 (14)	<.001
$M_{negative}$	2478.67 (347)	<.001	0.86	0.88	0.04	0.05	$M_{negative}$ vs. $M_{baseline}$	1211.35 (15)	<.001
$M_{2methods}$	2114.76 (333)	<.001	0.87	0.90	0.04	0.04	$M_{2methods}$ vs. $M_{baseline}$	1443.90 (29)	<.001
							$M_{2methods}$ vs. $M_{positive}$	4274.83 (15)	<.001
							$M_{2methods}$ vs. $M_{negative}$	292.22 (14)	<.001
M_{2m+r}	**2072.67 (332)**	**<.001**	**0.88**	**0.90**	**0.04**	**0.04**	M_{2m+r} vs. $M_{baseline}$	1497.16 (30)	<.001
							M_{2m+r} vs. $M_{positive}$	1076.12 (14)	<.001
							M_{2m+r} vs. $M_{negative}$	1211.35 (15)	<.001
							M_{2m+r} vs. $M_{2methods}$	289.90 (1)	<.001

Note: The final best model based on Bentler–Satorra (2001) scaled difference χ^2 tests is shown in bold. Because MLR was used for model estimators, we used the Satorra–Bentler (2001) test rather than the more common χ^2 difference test. The former test (Satorra–Bentler test) is appropriate when MLR is the estimator but the latter test (common χ^2 difference test) is appropriate when ML (maximum likelihood) is the estimator (Satorra & Bentler, 2001).
TLI = Tucker-Lewis Index; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual.

within a measure. The current study will examine whether acquiescence bias and item keying effect can exist simultaneously. This question is important because the estimation of one response style can sometimes disappear after another response style is controlled for. Second, we included multiple constructs in our study because response styles found in only one particular measure, as in Alessandri et al. (2010), may simply represent idiosyncratic characteristics of that measure. The HEXACO personality traits are used in the current study because they have been employed across multiple areas of psychology and each personality trait is theoretically unidimensional. If an item keying effect is found in these multiple unidimensional personality measures, it is likely to represent response patterns rather than representation of an unintended construct in a particular measure. Third, to our knowledge, the heritability of acquiescence bias has never been assessed in previous research. Our research is theoretically important because it assesses potential biological determinants in acquiescence bias. Finally, we will examine whether response styles affect heritability estimates; for example, does controlling for acquiescence bias affect the heritability estimates in twin studies in general? We will address the last research question by comparing the heritability estimates of personality constructs before and after controlling for response styles.

Method

Participants

Participants were 795 MZ pairs (726 female and 69 male pairs) and 662 same-sex DZ pairs (616 female and 46 male pairs). The average age of the twins was 58.56 years (SD = 12.85), and ranged in age from 18 to 86 years. Participants were part of a larger study (see Veselka et al., 2009) and were part of the Department of the TwinsUK registry based at the Department of Twin Research and Genetic Epidemiology (DTR), King's College, London, England.

Self-completion questionnaires were sent to these adult twins, who are all volunteers in the TwinsUK Adult Twin Registry (Spector & Williams, 2006). All were ascertained from the general population and shown to be comparable to age-matched population singletons. These unselected MZ and DZ twins have been recruited since 1992 using twin registers and national media campaigns and have been used in a wide variety of studies (www.twinsuk.ac.uk). For historical reasons, the cohort is predominantly female: when the twin study started, its purpose was to study bone diseases such as osteoporosis and osteoarthritis and the effects of hormones at menopause. After 2004, the study opened up to include males, but no targeted recruitment of males has been done. The twins in the registry are not selected for any particular trait and they volunteer to take part in studies that cover a wide range of traits and common medical conditions (Andrew et al., 2001). The study was approved by the St Thomas' Hospital research ethics committee, and all twins in the study provided informed consent.

Materials and Procedure

Participants completed the 60-item HEXACO Personality Inventory (HEXACO-60; Ashton & Lee, 2009) as part of a mailed survey package (see Veselka et al., 2009). The HEXACO assesses six dimensions of personality: honesty–humility, emotionality, extraversion, agreeableness, conscientiousness, and openness-to-experience. Items are responded by using a 1 (*strongly disagree*) to 5 (*strongly agree*) Likert scale. The internal consistency coefficients of the scales are shown in Table 1.

Results

Multitrait–Multimethod Confirmatory Factor Analyses

The method effect in the HEXACO personality scale was evaluated using the correlated trait–uncorrelated method (CTUM) model within a framework of

https://doi.org/10.1017/thg.2013.38 Published online by Cambridge University Press

multitrait–multimethod (MTMM) confirmatory factor analysis (CFA). CTUM was chosen over other MTMM CFA models (e.g., correlated trait–correlated method; CTCM) because it does not overestimate the method variance, thereby providing a more conservative estimation of the method effect (Marsh, 1989; see also Marsh & Bailey, 1991). Recent evidence by Marsh et al. (2010) supported the estimation accuracy of CTUM models. Following recommendations in the literature, we parceled personality items (e.g., Marsh et al., 1998; Nasser & Wisenbaker, 2003; Yuan et al., 1997). Parceled indicators are psychometrically more reliable than individual items. Items belonging to the same valence within a scale (e.g., positively valenced honesty items) were first factor-analyzed and then ordered according to their factor loadings. Based on a popular procedure, an item with high factor loadings is paired with another item with low factor loadings to form a parcel, so that the averaged loading becomes more balanced among the parcels (Yuan et al., 1997). Two to three items were averaged to form a parcel in the model. Finally, we followed a statistical procedure suggested by Stapleton (2006) in order to control for the statistical dependency within each pair of twins in our model estimations. This procedure, using a robust maximum likelihood estimator (MLR), provides more accurate estimation of the model parameters when the assumptions of independent observations and normality are violated.

We compared five nested MTMM CFA models (see Figure 1). The first model, M_{baseline}, is a common CFA model where observed personality item indicators load on their corresponding construct factors only. The second model, M_{positive}, builds upon the baseline model by including a positive valence method factor. All positively valenced items[1] load on this method factor. The third model, M_{negative}, includes a negative valence method factor and all negatively valenced items load on this method factor. The fourth model, $M_{2\text{methods}}$, includes both the positive and negative valence factors, whereas in the CTUM framework they are restricted to be orthogonally positioned to each other. If $M_{2\text{methods}}$ fits better than the previous three models (M_{baseline}, $M_{\text{positive,}}$ and M_{negative}), it is likely due to the fact that participants are showing differential response styles to items of opposing valence. The fifth model, M_{2m+r}, is built upon $M_{2\text{methods}}$ by including a random intercept (Maydeu-Olivares & Coffman, 2006) in addition to the two valence factors. Acquiescence bias is modeled by this random intercept (Maydeu-Olivares & Coffman, 2006).

The five MTMM models were estimated with the Mplus 6.1 computer program (Muthén & Muthén, 1998–2010; Table 1). Although M_{positive} and M_{negative} models fit better than M_{baseline}, the $M_{2\text{methods}}$ model showed better fit than any of these three models. This result replicates the finding by Biderman et al. (2011), suggesting that participants respond to positively valenced items and negatively valenced items differently. However, the best fitting model contains the positive and negative valence methods and a random intercept (M_{2m+r}) factor. Results suggest that both acquiescence bias and the two valence factors exist in participants' responses. We also decomposed the variance explained by the personality factors, acquiescence bias, item keying response style, and residual variance. The results are as follows: 33.18% (personality factors), 5.73% (item keying), 2.11% (acquiescence bias), and 58.98% (residual variance). We then selected the best-fitting MTMM model (M_{2m+r}) and extracted the factor scores from the method factor(s) and the six HEXACO personality factors from the chosen model for heritability analyses.

Univariate Genetic Analyses

Analysis of the factor score residual covariance (controlling for sex and age), using standard SEM methods (Neale & Cardon, 1992), was performed (Neale et al., 2006). In conducting univariate genetic analyses, a phenotypic score is expressed as a linear function of three factors: genetic (A), common environment (C), and specific environment (E). Following the finding of Sullivan and Eaves (2002), a full ACE model was analyzed and not the reduced models. Table 2 presents the intraclass correlations for the MZ and DZ twins for each factor score as well as the estimated heritability (a^2), common environment (c^2), and specific environment (e^2) values from the standardized parameter estimates, and the 95% confidence interval (intervals not containing zero are deemed to be significant). All of the factor scores, except for the acquiescence bias, were found to have a significant genetic component, with values ranging from 0.24 for the positive valence factor to 0.54 for openness to experience. Common environmental effects were found to account for zero percent of the variance for personality factors and item keying factors. Unique environmental effects were significant for all of the factor scores and ranged from 0.46 for openness to experience to 0.76 for the positive valence factor. Our analysis thus suggested that both positive and negative item keying factors are heritable but acquiescence bias is not. Neither acquiescence bias nor item keying effect is found to be significantly influenced by the environment shared between the twins.

Finally, we examined whether controlling for the two response styles would substantially affect the heritability estimates of HEXACO personality factors. We compared estimates of the personality factors between the baseline model (M_{baseline}; Table 3) and our final selected model (M_{2m+r}; Table 2). The estimates do not differ substantially. The extraction of method factors in M_{2m+r} does not substantially change the heritability estimates of personality traits.

Discussion

The current study advanced previous research in at least three significant ways. First, we were able to show both acquiescence bias and item keying response styles in our personality data. Previous researchers often examined and statistically controlled for only one of these response styles

https://doi.org/10.1017/thg.2013.38 Published online by Cambridge University Press

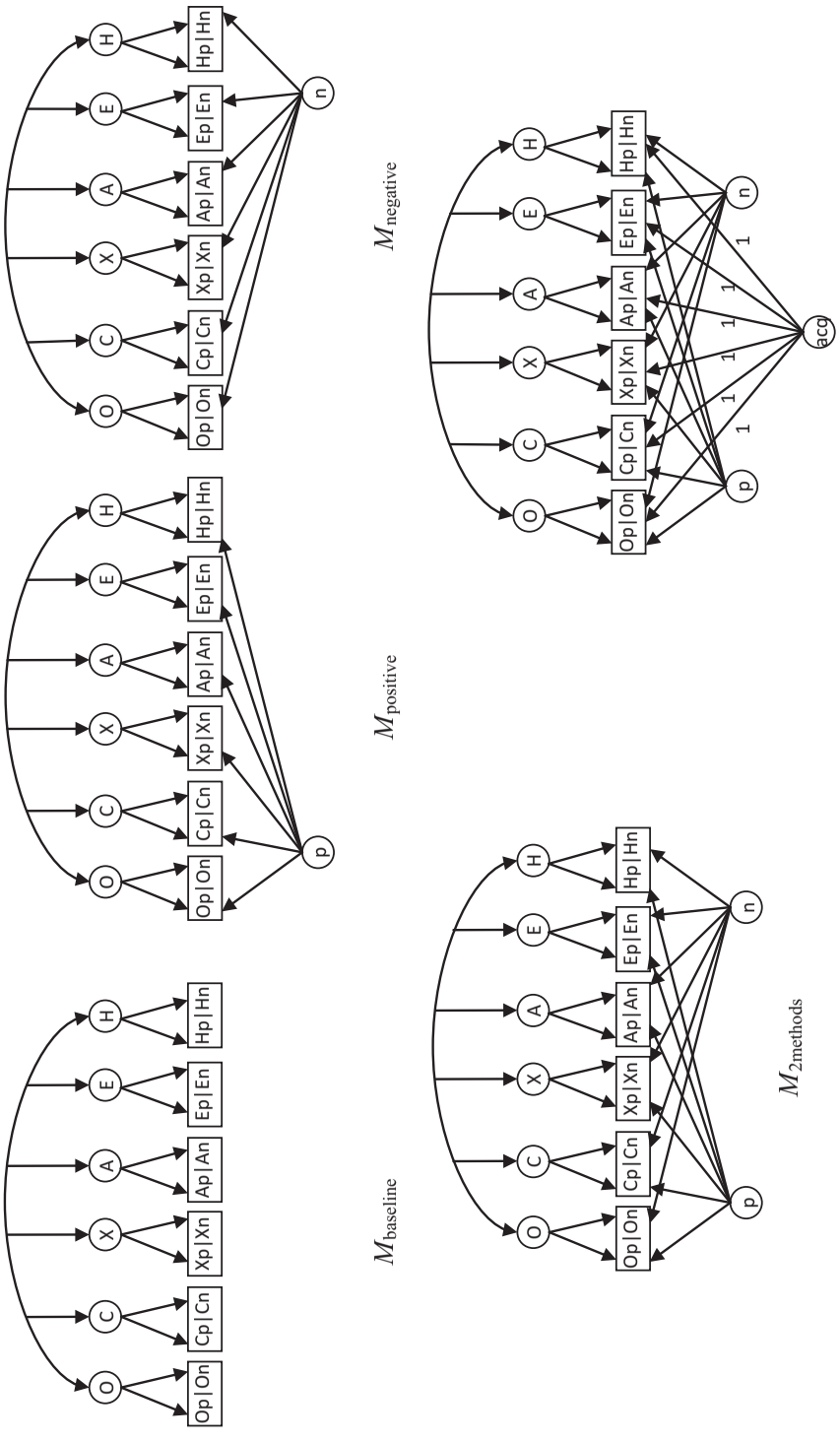

FIGURE 1

MTMM models. Note: O = openness; C = conscientiousness; X = extraversion; A = agreeableness; E = emotionality; H = honesty–humility; p = positive valence factor; n = negative valence factor; acq = acquiescence bias. Each rectangle represents observed indicators that are measured by positively valenced items (e.g., Op) and negatively valenced items (e.g., On). Single arrows represent factor loadings and double arrows represent covariance among latent factors.

https://doi.org/10.1017/thg.2013.38 Published online by Cambridge University Press

TABLE 2

Scale Internal Consistency and Univariate Genetic Results for Final Model (M_{2m+r})

	Cronbach's α	MZr_i	DZr_i	a^2	c^2	e^2
Honesty	0.66	0.31	0.13	0.30 (0.16 to 0.36)	0.00 (0.00 to 0.11)	0.70 (0.64 to 0.76)
Emotionality	0.73	0.47	0.16	0.45 (0.35 to 0.50)	0.00 (0.00 to 0.07)	0.55 (0.50 to 0.61)
Extraversion	0.76	0.46	0.14	0.43 (0.34 to 0.48)	0.00 (0.00 to 0.07)	0.57 (0.52 to 0.62)
Agreeableness	0.74	0.33	0.10	0.31 (0.20 to 0.37)	0.00 (0.00 to 0.08)	0.69 (0.63 to 0.75)
Conscientiousness	0.71	0.34	0.14	0.33 (0.19 to 0.38)	0.00 (0.00 to 0.11)	0.67 (0.62 to 0.73)
Openness	0.75	0.53	0.25	0.54 (0.43 to 0.59)	0.00 (0.00 to 0.09)	0.46 (0.41 to 0.50)
Pos	-	0.26	0.08	0.24 (0.10 to 0.30)	0.00 (0.00 to 0.11)	0.76 (0.70 to 0.82)
Neg	-	0.34	0.15	0.33 (0.15 to 0.38)	0.00 (0.00 to 0.14)	0.67 (0.62 to 0.73)
Acq	-	0.26	0.17	0.16 (0.00 to 0.31)	0.09 (0.00 to 0.24)	0.75 (0.68 to 0.81)

Note: r = intraclass correlation; pos = positive valence factor; neg = negative valence factor; acq = acquiescence bias; MZ = monozygotic; DZ = dizygotic.
Values in parentheses represent the 95% confidence interval.

TABLE 3

Univariate Genetic Results for Baseline Model ($M_{baseline}$)

	MZr_i	DZr_i	a^2	c^2	e^2
Honesty	0.30	0.15	0.30 (0.09 to 0.36)	0.00 (0.00 to 0.17)	0.70 (0.63 to 0.77)
Emotionality	0.48	0.14	0.45 (0.37 to 0.51)	0.00 (0.00 to 0.06)	0.55 (0.49 to 0.60)
Extraversion	0.50	0.09	0.45 (0.37 to 0.50)	0.00 (0.00 to 0.06)	0.55 (0.50 to 0.61)
Agreeableness	0.37	0.14	0.35 (0.20 to 0.41)	0.00 (0.00 to 0.12)	0.65 (0.59 to 0.71)
Conscientiousness	0.32	0.13	0.31 (0.16 to 0.37)	0.00 (0.00 to 0.12)	0.69 (0.63 to 0.75)
Openness	0.50	0.22	0.50 (0.37 to 0.55)	0.00 (0.00 to 0.11)	0.50 (0.45 to 0.56)

Note: r = intraclass correlation.
Values in parentheses represent the 95% confidence interval.

(e.g., Billiet & McClendon, 2000; Marsh et al., 2010; Motl & DiStefano, 2002; Quilty et al., 2006), and rarely attempted to model both biases simultaneously. An important advantage of using MTMM SEM techniques is that it allows us to model bias within the same framework and examine whether one type of bias can be fully explained by another bias. Our findings, however, suggested that both types of bias must be modeled to achieve the best fit to the data. In this way, our findings suggest future researchers should statistically control for both types of bias in their research, even though these response biases may not affect heritability estimates.

Second, an important advantage of modeling both response styles simultaneously is that it allows us to compare the magnitude between acquiescence bias and item keying, a comparison that is rarely considered. Although we found both acquiescence bias and item keying response bias within the dataset, their magnitude of influence on participants' responses differed substantially. Item keying response styles explained 5.90% of the variance while acquiescence explained 2.11%. This result suggests that item keying affects variance in participants' responses more than twice that of acquiescence bias in our sample. Although some researchers have suggested that response styles may not substantially affect the validity of a construct score (Schimmack et al., 2002), the current research does not support this position. Variance explained by the two types of response bias (7.84%) is over one-fifth of the variance explained by personality constructs (33.18%). Consequently, response styles are substantive parts of construct measurement.

Third, most importantly, we found that the item keying effect in personality is heritable whereas acquiescence bias in personality is not. Past research by Alessandri et al. (2010) only tested the heritability of item keying effect in relation to one construct (i.e., optimism). As we have explained, the reverse-keyed (negatively valenced) items of their optimism scale may actually represent the measurement of another construct (pessimism). However, with the personality constructs from HEXACO, ambiguity in the item keying effect is mitigated. Each personality dimension included in the current study is theoretically a unidimensional construct (Paunonen & Hong, in press; Wiggins, 1973). In addition, we have included multiple measures in the study. For these reasons, the response styles detailed in the current study are unlikely to represent an opposite pole of any one particular personality construct, but rather a response bias on oppositely valenced personality items in general. Our methodology then overcomes a major limitation of Alessandri et al.'s (2010) study, which relies on only one construct in their investigation.

Perhaps due to the methodological improvement of the current study, our results differ substantially from Alessandri et al. (2010), who found that only the positive valence factor in an optimism measure is heritable. When we examine positive and negative valence effects in univariate twin analyses, we find that both of these factors are heritable. In addition, the heritability was 24% for the positive valence factor and 33% for the negative valence factor, and both factors were not influenced by the environment that is shared by the twins. As compared with the heritability of

https://doi.org/10.1017/thg.2013.38 Published online by Cambridge University Press

personality traits in the current study (30–54%), our results thus suggest that the heritability of positive and negative valence factors is at the lower end of personality traits. This finding regarding the heritability of item keying is extremely important because it implies that participants' responses to valence represent something that would be temporally stable and thus probably meaningful as it may be related with human biological system.

Contrary to the item keying response styles, our results revealed that acquiescence bias has little or no heritability. When we compared the correlation of acquiescence bias between MZ twins and DZ twins, we find a stronger relation for the former group. In spite of this, the univariate genetic analyses failed to find statistical significance in its heritability estimate. This result thus suggests the possibility that acquiescence bias is at least not as heritable as personality factors and item keying effects. The only significant influence on acquiescence bias was environmental factors that are not uniquely shared by the twins. This finding is interesting and warrants further investigation.

As with previous researchers (Alessandri et al., 2010), we are still uncertain why item keying response styles are influenced primarily by heredity rather than shared environment (DiStefano & Motl, 2006; Quilty et al., 2006; Rauch et al., 2007). In contrast, past research on social and cultural psychology may already provide answers to the non-heritability of acquiescence bias. Cross-cultural psychologists have consistently showed that acquiescence is amenable to influences of external cultural factors. For example, there is overwhelming evidence that respondents from collectivistic cultures (e.g., East Asians) are more likely to acquiesce as compared to respondents from individualistic cultures (e.g., North Americans; Chen et al., 1995; Johnson et al., 2005; Smith, 2004; van Hemert et al., 2002). Collectivistic cultures encourage individuals to fit into an existing social structure and to conform to the overall goals or wishes of other members in the society. This experience affects their communication styles — they are more likely to acquiesce because doing so promotes relationship harmony and avoids conflict with others (Johnson et al., 2005; see also Kwan et al., 1997; Peng & Nisbett, 1996; Smith & Fischer, 2008 for a discourse on this topic). If we look at our data closely, the results show the *possibility* that acquiescence is more amenable to shared environment between the twins than do other personality factors ($c^2 = 0.09$ vs. 0.00), although one may caution not to overinterpret this finding as its confidence interval for acquiescence overlaps with those for other personality factors. Given that our twin sample is mature in their age (90% of the twins are between the age of 34 and 77; mean age = 58.56 years), there are plenty of unshared, external environmental influences (i.e., the 'E' component in the model) on the twins once they live apart. The non-significance of the 'C' component for acquiescence may suggest that this bias is amenable to external influences even later in life, just as a person can still be adapted to cultures at least after adolescence (Oyserman & Lee, 2008). The non-significance of heritability (i.e., the 'A' component) implies that biological factors are not central to the development of this response style.

Limitations and Future Directions

As with any other empirical studies, the current study has several limitations. First, our sample represents older twins from Britain and little is known regarding the generalizability of our findings to younger twins. Future research may try to replicate our results with younger twins or with people from other cultural backgrounds. Second, the sample is predominantly female so it would also be important to replicate the results with a more evenly sex-balanced sample. Third, although we included constructs from a commonly used personality model (HEXACO), we do not know whether item keying response styles from constructs outside personality are heritable. Therefore, confirmation of these results with other psychological constructs will be valuable. Finally, the current study investigates participants' response styles to a 5-point Likert scale. Little is known whether the result will be generalizable to other response formats such as a 7-point Likert scale format or a true–false format. Although we do not expect the results will differ based on the scaling structure, replications of our findings in these other formats will be important.

Our study suggests researchers should shift their thinking regarding acquiescence bias and item keying response styles as representing something unworthy of empirical study. The fact that the heritability estimates of item keying response styles are at a similar level of some personality traits is itself intriguing and is worthy of further exploration. Acquiescence bias, although not as heritable as item keying response styles or personality factors, was found to be largely shaped by unshared environmental influences. Future research should continue to explore the nature of item keying response styles and acquiescence bias, because they are potentially important psychological phenomena that inform our understanding of human cognition and behaviors.

Acknowledgments

We would like to thank the twins for their voluntary contribution to this research project. We would also like to thank the staff of the Department of Twin Research for their help and support in undertaking this project. The Wellcome Trust provides core support for DTR. The study also receives support from the Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London. Finally, we thank John Meyer for his valuable feedback on an earlier version of the manuscript.

https://doi.org/10.1017/thg.2013.38 Published online by Cambridge University Press

Endnote

1 Positive-valenced items are the regular-keyed items of honesty–humility, extraversion, agreeableness, conscientiousness, and openness, and reverse-keyed items of emotionality. Negative-valenced items are the reverse-keyed items of honesty–humility, extraversion, agreeableness, conscientiousness, and openness, and regular-keyed items of emotionality.

References

Alessandri, G., Vecchione, M., Fagnani, C., Bentler, P. M., Barbaranelli, C., Medda, E., . . . Caprara, G. V. (2010). Much more than model fitting? Evidence for the heritability of method effect associated with positively worded items of the life orientation test revised. *Structural Equation Modeling*, *17*, 642–653.

Andrew, T., Hart, D. J., Snieder, H., De Lange, M., Spector, T. D., & MacGregor, A. J. (2001). Are twins and singletons comparable? A study of disease-related and lifestyle characteristics in adult women. *Twin Research and Human Genetics*, *4*, 464–77.

Ashton, M. C., & Lee, K. (2009). The HEXACO-60: A short measure of the major dimensions of personality. *Journal of Personality Assessment*, *91*, 340–345.

Ayidiya, S. A., & McClendon, M. J. (1990). Response effects in mail surveys. *Public Opinion Quarterly*, *54*, 229–247.

Bachman, J. G., & O'Malley, P. M. (1984). Black–white differences in self-esteem: Are they affected by response styles? *American Journal of Sociology*, *90*, 624–639.

Biderman, M. D., Nguyen, N. T., Cunningham, C. J. L., & Ghorbani, N. (2011). The ubiquity of common method variance: The case of the Big Five. *Journal of Research in Personality*, *45*, 417–429.

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 and 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.

Block, J. (1965). *The challenge of response sets*. New York: Century.

Chen, C., Shin-ying, L., & Stevenson, H. W. (1995). Response style and cross-cultural comparisons of rating scales among East Asian and North American students. *Psychological Science*, *6*, 170–175.

Couch, A., & Keniston, K. (1960). Yessayers and naysayers: Agreeing response set as a personality variable. *Journal of Abnormal and Social Psychology*, *60*, 151–174.

Cronbach, L. J. (1941). An experimental comparison of the multiple true-false and multiple multiple-choice tests. *Journal of Educational Psychology*, *32*, 533–543.

Cronbach, L. J. (1942). Studies of acquiescence as a factor in the true-false test. *Journal of Educational Psychology*, *33*, 401–415.

DiStefano, C., & Motl, R. W. (2006). Further investigating method effects associated with negatively worded items on self-report surveys. *Structural Equation Modeling*, *13*, 440–464.

Grimm, S. D., & Church, A. T. (1999). A cross-cultural study of response biases in personality measures. *Journal of Research in Personality*, *33*, 415–441.

Heal, L. W., & Sigelman, C. K. (1995). Response biases in interviews of individuals with limited mental ability. *Journal of Intellectual Disability Research*, *39*, 331–340.

Jackson, D. N., & Messick, S. (1958). Content and style in personality assessment. *Psychological Bulletin*, *55*, 243–252.

Johnson, T., Kulesa, P., Cho, Y. I., & Shavitt, S. (2005). The relation between culture and response styles: Evidence from 19 countries. *Journal of Cross-Cultural Psychology*, *36*, 264–277.

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.

Kam, C., Zhou, X., Zhang, X., & Ho, M. Y. (2012). Examining the dimensionality of self-construals and individualistic-collectivistic values with random intercept item factor analysis. *Personality and Individual Differences*, *53*, 727–733.

Krosnick, J. A., Narayan, S., & Smith, W. R. (1996). Satisficing in surveys: Initial evidence. In M. T. Braverman & J. K. Slater (Eds.), *Advances in survey research* (pp. 29–44). San Francisco: Jossey-Bass.

Kwan, V. S. Y., Bond, M. H., & Singelis, T. M. (1997). Pancultural explanations to life satisfaction: Adding relationship harmony to self-esteem. *Journal of Personality and Social Psychology*, *5*, 1038–1051.

Lenski, G. E., & Leggett, J. C. (1960). Caste, class, and deference in the research interview. *American Journal of Sociology*, *65*, 463–467.

Lentz, T. F. (1938). Acquiescence as a factor in the measurement of personality. *Psychological Bulletin*, *35*, 659.

Lorge, I. (1937). Gen-like: Halo or reality? *Psychological Bulletin*, *34*, 545–546.

Marsh, H. W. (1989). Confirmatory factor analyses of multitrait–multimethod data: Many problems and a few solutions. *Applied Psychological Measurement*, *13*, 335–361.

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

Marsh, H. W., & Bailey, M. (1991). Confirmatory factor analysis of multitrait–multimethod data: A comparison of the behavior of alternative models. *Applied Psychological Measurement*, *15*, 47–70.

Marsh, H. W., Hau, K. T., Balla, J. R., & Grayson, D. (1998). Is more ever too much? The number of indicators per factor in confirmatory factor analysis. *Multivariate Behavioral Research*, *33*, 181–220.

Marsh, H. W., Scalas, L. F., & Nagengast, B. (2010). Longitudinal tests of competing factor structures for the Rosenberg Self-Esteem scale: Traits, ephemeral artifacts, and stable response styles. *Psychological Assessment*, *22*, 366–381.

https://doi.org/10.1017/thg.2013.38 Published online by Cambridge University Press

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

Meisenberg, G., & Williams, A. (2008). Are acquiescent and extreme response styles related to low intelligence and education? *Personality and Individual Differences*, *44*, 1539–1550.

Messick, S., & Frederiksen, N. (1958). Ability, acquiescence, and "authoritarianism". *Psychological Reports*, *4*, 687–697.

Messick, S., & Frederiksen, N. (1961). Acquiescence and desirability as response determinants on the MMPI. *Educational and Psychological Measurement*, *21*, 771–790.

Messick, S., & Jackson, D. N. (1961). Acquiescence and the factorial interpretation of the MMPI. *Psychological Bulletin*, *58*, 299–304.

Motl, R. W., & DiStefano, C. (2002). Longitudinal invariance of self-esteem and method effects associated with negatively worded items. *Structural Equation Modeling*, *9*, 562–578.

Muthén, L. K., & Muthén, B. O. (1998–2010). *Mplus user's guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.

Narayan, S., & Krosnick, J. A. (1996). Education moderates some response effects in attitude measurement. *Public Opinion Quarterly*, *60*, 58–88.

Nasser, F., & Wisenbaker, J. (2003). A Monte Carlo study investigating the impact of item parceling on measures of fit in confirmatory factor analysis. *Educational and Psychological Measurement*, *63*, 729–757.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2006). *Mx: Statistical modeling manual* (7th ed.). Richmond, VA: Department of Psychiatry, Medical College of Virginia.

Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. Norwell, MA: Kluwer.

Oyserman, D., & Lee, S. W. S. (2008). Does culture influence what and how we think? Effects of priming individualism and collectivism. *Psychological Bulletin*, *134*, 311–342.

Paunonen, S. V., & Hong, R. Y. (in press). In defense of personality traits. In P. R. Shaver & M. Mikulincer (Eds.), *Handbook of personality and social psychology*. New York: APA.

Peabody, D. (1967). Trait inferences: Evaluative and descriptive aspects. *Journal of Personality and Social Psychology Monograph*, *7*(4, Serial No. 644), 1–18.

Peng, K., & Nisbett, R. (1996). Cross-cultural similarities and differences in understanding of physical causality. In G. Shield & M. Shale (Eds.), *Culture and science* (pp. 10–21). Frankfort, KY: Kentucky State University Press.

Quilty, L. C., Oakman, J. M., & Risko, E. (2006). Correlates of the Rosenberg self-esteem scale method effects. *Structural Equation Modeling*, *13*, 99–117.

Rammstedt, B., Goldberg, L. R., & Borg, I. (2010). The measurement equivalence of Big-Five factor markers for persons with different levels of education. *Journal of Research in Personality*, *44*, 53–61.

Rauch, W. A., Moosbrugger, H., & Schweizer, K. (2007). Method effects due to social desirability as a parsimonious explanation of the deviation from unidimensionality in LOT–R scores. *Personality and Individual Differences*, *42*, 1597–1607.

Rorer, L. G. (1965). The great response-style myth. *Psychological Bulletin*, *63*, 129–156.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Satorra, A., & Bentler, P.M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507–514.

Schimmack, U., Böckenholt, U., & Reisenzein, R. (2002). Response styles in affect ratings: Making a mountain out of a molehill. *Journal of Personality Assessment*, *78*, 461–483.

Sigelman, C. K., Budd, E. C., Spanhel, C. L., & Schoenrock, C. J. (1981). When in doubt, say yes: Acquiescence in interviews with mentally retarded persons. *Mental Retardation*, *19*, 53–58.

Smith, P. B. (2004). Acquiescent response bias as an aspect of cultural communication style. *Journal of Cross-Cultural Psychology*, *35*, 50–61.

Smith, P. B., & Fischer, R. (2008). Acquiescence, extreme response bias and levels of cross-cultural analysis. In F. J. R. van de Vijver, D. A. van Hemert & Y. Poortinga (Eds.), *Individuals and cultures in multi-level analysis* (pp. 283–311). Mahwah, NJ: Lawrence Erlbaum Associates.

Spector, T. D., & Williams, F. M. (2006). The UK Adult Twin Registry (TwinsUK). *Twin Research and Human Genetics*, *9*, 899–906.

Stapleton, L. M. (2006). An assessment of practical solutions for structural equation modeling with complex sample data. *Structural Equation Modeling*, *13*, 28–58.

Sullivan, P. F., & Eaves, L. J. (2002). Evaluation of analyses of univariate discrete twin data. *Behavior Genetics*, *32*, 221–227.

Tomás, J. M., & Oliver, A. (1999) Rosenberg's self-esteem scale: Two factors or method effects. *Structural Equation Modeling*, *6*, 84–98.

van Hemert, D. A., van de Vijver, F. J. R., Poortinga, Y. H., & Georgas, J. (2002). Structural and functional equivalence of the Eysenck Personality Questionnaire within and between countries. *Personality and Individual Differences*, *33*, 1229–1249.

Veselka, L., Schermer, J. A., Petrides, K. V., Cherkas, L. F., Spector, T. D., & Vernon, P. A. (2009). A general factor of personality: Evidence from the HEXACO model and a measure of trait emotional intelligence. *Twin Research and Human Genetics*, *12*, 420–424.

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

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

Yuan, K.-H., Bentler, P. M., & Kano, Y. (1997). On averaging variables in a CFA model. *Behaviormetrika*, *24*, 71–83.

https://doi.org/10.1017/thg.2013.38 Published online by Cambridge University Press