Beyond Questionable Research Methods: The Role of Intellectual Honesty in Research Credibility

Frank L. Schmidt

University of Iowa

Author Note: Frank L. Schmidt, Department of Management and Organizations, University of Iowa, Iowa City, IA 52242. Enquiries concerning this paper should be sent to Frank L. Schmidt, Department of Management and Organizations, University of Iowa, Iowa City 52242. Email: frank-schmidt@uiowa.edu

Beyond Questionable Research Methods: The Role of Intellectual Honesty in Research Credibility

Abstract

This paper explores an important credibility problem in the research literature beyond the issue of questionable statistical methods in research: the problem of apparent lack of intellectual honesty in many published research articles. I focus on this problem in two areas: (a) studies purporting to demonstrate the effects of people's experiences on their later life outcomes while failing to discuss or mention the probable causal role of genetic inheritance in producing these effects, despite the overwhelming evidence for this connection from behavior genetics research; and (b) studies of specific aptitudes (specific abilities) such as verbal, spatial, or reasoning that fail to acknowledge or mention that such aptitudes are indicator variables for general mental ability (GMA; or intelligence) and that after proper control for GMA the residuals in these aptitudes make no contribution to prediction of real world academic, occupational, or job performance. It is only the GMA component in such aptitudes that produces the ability to predict. As is well known today, the issue of the credibility of research conclusions is prominent (Ioannidis, 2005). In both the areas examined in this paper, these deficiencies create serious and unnecessary credibility problems, and the doubts they inspire about credibility could unfortunately be generalized to other research areas in which these problems do not exist.

Keywords: behavior genetics, heritability, general mental ability, Intelligence, specific aptitudes, specific abilities, research credibility, intellectual honesty.

In recent years there have been numerous published critiques of the credibility of research findings in psychology (and in other fields: medical, pharmacology, economics, and others.). These critiques have focused on questionable statistical research methods and practices. (This includes one article by the present author and a co-author; Schmidt and Oh, 2016). However, many conclusions presented in research studies are questionable not because of the data analysis methods used, but because of a certain lack of intellectual honesty, namely failure to consider, discuss, or even mention well-established research findings that are highly relevant to the article's content and conclusions. I must assume that this lack of intellectual honesty is typically inadvertent and unintended, stemming from lack of knowledge of the relevant research findings. This paper examines two areas of research in which this is a major problem: (a) research on the effects of earlier experiences or exposures on later life outcomes; and (b) research on specific abilities (aptitudes). In both these areas, this failure threatens the credibility of the research.

Failure to Acknowledge Well Established Behavior Genetics Findings

The first area of problem research focuses on the ostensible effects of life experiences on life outcomes. The aspect of much of this research that is lacking in intellectual honesty is the common failure to acknowledge the relevant findings in the field of behavior genetics. These findings show that virtually all tendencies, traits, behaviors, and life outcomes have a substantial genetic basis (c.f., Bouchard, 1997a, 1997b, 2004; Colarelli & Arvery, 2015; Lee & McGue, 2016; McGue & Bouchard, 1998; Plomin, DeFries, Knopik, & Neiderhaise, 2013; Plomin, DeFries, Knopik, & Neiderhaise., 2016; Plomin, Owen, & McGuffin, 1994; Turkheimer, 2000).

Even day-to-day variability in positive and negative affect has been shown to be substantially heritable (Zheng, Plomin, & von Stumm, 2016).

Research has further shown that most supposedly purely environmental variables (such as the number of books and magazines in the home) that are often concluded to be causes of later life outcomes are themselves genetically influenced (e.g., see Plomin & Bergman, 1991; Plomin et al., 2016). That is, they are substantially influenced by the genetic makeup of the parents in the home, whose genes are passed on to their offspring. Research also indicates that people seek out and create their own environments based on their genetically influenced proclivities and interests (Scar, 1996; Scar, 1989; Scar & McCartney, 1983).

These behavior genetics findings do not mean that experiences of people do not have any effect on their later life outcomes. But they do mean that failure to even mention potential or likely genetic effects on these outcomes is a serious problem, one that reduces the credibility of the research. We now examine some examples of studies that fail to acknowledge these well-established research findings.

A number of studies report that children who grow up in dysfunctional or abusive families tend later as adults to be abusive themselves (cf. Kaufman & Zigler, 1987; 1989). The interpretation is virtually always entirely environmental: It is assumed that the earlier experiences cause the later behavior. There is no acknowledgement of the fact that all major behaviors, including abusive personalities, have a genetic component (Plomin, Owen, & McGuffin, 1994; Plomin et al., 2016; Turkheimer, 2000). There is no mention of the possibility that the genes that lead the parents to be abusive are passed on to their children and are an important reason why their

children later themselves also become abusive as adults (Rowe, 1994). There are numerous examples of such studies in the literature (Kaufman & Zigler, 1987; 1989).

Warlaumont, Richards, Gilkerson, & Oller (2014) found that the parents of autistic children talk less to their children than other parents do; the article concludes that this is a cause (maybe *the* cause) of autism. There is no mention of genetics, despite the fact that autism has been shown to be strongly genetic (Lord, Cook, Leventhal, & Amaral, 2000; Sigman & Capps, 1997; Plomin et al., 2013). It is highly likely that the reduced amount of talking to children among parents of autistic children is caused by one or both of the parents being at least somewhat on the autism spectrum themselves and that their children are autistic for reasons of genetic inheritance. Stoltenberg and Burmeister (2000) point out that some behavioral abnormalities are present in the parents of autistic children and that these abnormalities "may be markers for some of the many predisposing genes for autism." Yet the Warlaumont et al. article contains no mention of any possible genetic connection.

Taylor, Manganello, Lee, and Rice (2010) found that 3-year old children who were spanked by their mothers showed more aggressive behavior at age 5 than those children who had not been spanked. They concluded that the earlier spanking caused the later aggressive behavior. There was no mention of the possibility that the personality traits of the children were the causal variables at both time periods. The possibility that it is the most difficult, misbehaving, and aggressive children who are spanked and these difficult behaviors and the personality traits underlying them are stable and continue to manifest themselves at age 5 was pointed out by Baumrind (2009). As noted above, personality traits have a genetic basis.

Waldinger and Schulz (2016) found that people who experienced nurturing family environments when young showed better emotion regulation styles as adults and were more secure in intimate relationships. In this article, the relations observed are interpreted as being completely due to environmental effects. There is no acknowledgment of the likelihood that the same genetic effects that cause the parents to generate a nurturing home environment are passed on to their children, causing them to show better adjustment later in life.

The study by Lippold, Davis, McHale, and Buxton (2016) is somewhat similar to the Waldinger and Schulz study. Lippold et al. examined the reactions of adolescents to daily stressors. They found that youth who experienced more parental warmth showed lower negative affect in response to stressors and also a faster rate of blood cortisol decline after experiencing a stressor. The interpretation was that parental warmth was the causal variable. There is no mention of the possibility that both the warmth of the parents and the stress reactions of the offspring could have the same genetic basis, a definite possibility in light of the behavior genetics research findings. Genetically loaded personality traits shared by parents and offspring could well be the causative link.

Several studies reported that children who are exposed to the classical music of Mozart have higher IQs (Campbell, 2001; Hetland (2000). The implication was that the music caused increased IQs. There is no mention in these studies of the possibility that parental general mental ability (GMA) and socio-economic status (SES) are the real causes. Classical music is more often played in higher SES homes and parental GMA levels average higher in such homes. And GMA is highly heritable (.75; Bouchard, 1997b). When experimental studies were conducted that controlled for this, the Mozart Effect was found to be non-existent (Steele,

Bass, & Crook, 1999). Nevertheless, these reports led Governor Zell Miller of North Carolina to arrange for all families with children in that state to receive a CD of Mozart's music (Author, 2010; Mackenzie, 1999).

Scar (1996) notes that several studies report that *authoritative* parenting produces well-adjusted children, while *authoritarian* parenting results in children who are more prone to behavior problems and have lower school achievement. In these studies, there is no discussion of the likelihood that parental GMA is the cause of this relation. Higher GMA parents are more likely to employ authoritative parenting styles and lower GMA parents are more likely to employ authoritarian parenting styles (Scar, 1996). Both GMA (Bouchard, 1997b) and Authoritarianism (Waller, Kojetin, Bouchard, Lykken, & Telligen, 1990) are genetically loaded, and children can inherit both from their parents.

As discussed by Scar (1989), some studies report that children of parents who read to them have higher IQs (GMA) and better school achievement. The implication is that the reading is the cause of the higher IQs. In fact, this linkage is substantially genetic in nature. Higher GMA parents are more likely to read to their children, and their children are more likely to inherit genes that contribute to higher GMA. Scar (1989) showed that the strongest link with the child's IQ was the average IQ of the parents. After controlling for this there was essentially no relation between the child's IQ and any of a number of parental characteristics or child rearing practices (e.g., parental knowledge of child development; parental education; mother's positive discipline; mother's positive control). Other studies that fail to recognize the genetically causative role of parental GMA on abilities, characteristics, and behaviors of their children include: Brown, Mounts, Lamborn & Steinberg (1993); Darling & Steinberg (1992); DeBaryshe,

Patterson, & Capaldi (1993); Durbin, Darling, Steinberg, & Brown (1993); Tang, Davis-Kean, Sexton, & Chen (2014); and Wachs & Gruen (1982).

An article by Chen and Miller (2012) attempts to explain why some children growing up in extremely disadvantaged circumstances nevertheless do well and even thrive. The explanation they present is entirely in terms of environmental effects, such as the influence of mentors on their development of strategies for dealing with adversity. There is no mention of potentially relevant personality traits, such as extroversion or conscientiousness, which are substantially heritable (.50; Bouchard, 1997b). Nor is there any discussion of the potential role of GMA, which is even more heritable (.75; Bouchard, 1997b). In fact, there is no mention of any individual traits the might contribute to resilience.

Bank, Burraston, and Snyder (2004) report that coercive parenting and frequent conflicts with siblings are linked to anti-social behavior and peer difficulties in children. The interpretation is that these effects are caused by the children's experience of coercive or ineffective parenting. There is no mention of the likely role of genetics. The personality traits of Disagreeableness, Neuroticism, and Authoritarianism are quite heritable (heritabilities are about .50; Bouchard, 1997b). Parents who are coercive are likely to be high on one or more of these traits and their children can inherit these same tendencies, leading them to engage in more conflict with their siblings and to show more anti-social behaviors. Another study manifesting this same problem is MacKinnon, Starnes, Volling, and Johnson (1997).

An article by Chen, Turiano, Mroczek et al. (2016) found that women who retrospectively report parental physical or emotional abuse in childhood have a higher incidence of all-cause

mortality as adults. The implication is that the cause of the elevated morality is the experience decades earlier of the abuse. Their article contains no mention of the possibility that the relation is trait-based or genetic in origin. The tendency towards angry outbursts, violence, and abuse behavior, like all such behaviors, is in part genetically based and can be inherited. It is also possible that this sort of personality leads to behaviors and conflicts that harm health and promote higher adult mortality rates.

A study by Daly, Quigley, Delaney, & Baumeister (2016), based on 21,132 individuals, found that self-control in childhood (at age 10) predicted ability to refrain from smoking in adulthood. This study included a partial control for cognitive ability but no control for potentially relevant personality traits (which, as noted earlier, have substantial genetic components). For example, there was no mention that the causal variable here could have been the personality trait of Conscientiousness, which could produce both childhood self-control and later ability to resist the temptation to smoke (self-control in adulthood).

The subjects in Davis, Carlo, Schwartz et al. (2016) were Latino adolescents who were recent immigrants to the U.S. The study found that those who reported feelings of being discriminated against showed more depressive symptoms and displayed fewer prosocial behaviors. The interpretation was that the “experience of discrimination” causes depression and reduces prosocial behaviors. There was no mention of the possible role of the genetically based personality trait of Neuroticism. The anxiety and other negative emotions characteristic of Neuroticism could be the cause of both the perceptions of discrimination and the depressive symptoms, with the symptoms of depression being the cause of the reduction in pro-social

behaviors. Neuroticism, like other personality traits, has been shown to have a substantial genetic component.

The social consequences of this form of lack of intellectual honesty in research are serious. One consequence is widespread public belief in false causal connections. For example, consider the finding that people who drink wine are healthier and live longer than people who drink other alcoholic beverages. The proposition expressed or implied in such articles is that drinking wine causes health and longevity. The real cause is likely to be socio-economic status (SES). Higher SES people are more likely to prefer wine. They also have better health in general and better health care, both self-care and medical care. And SES is itself partially genetically determined (Belsky, Moffitt, et al., 2016; Judge, Klinger, & Simon, 2010; Plomin et al., 2013). Bouchard (1997b) reports an average heritability of .45 for SES (occupational status). SES is also substantially correlated with GMA, which has an even larger genetic loading. There are other such examples, many of which are illustrated in the studies described above. Widespread false causal beliefs in the general population can cause many problems, especially in an active democracy in which many organizations advocate and implement interventions intended to improve human welfare. Wasted spending on intervention programs that have no causal effect may be common. For example, a campaign to get people to switch from other alcoholic beverages to wine.

Another social consequence of this form intellectual dishonesty is widespread overestimation of the potential of interventions to reduce or eliminate individual differences in important traits or behaviors (intelligence, academic achievement, etc.). The goal of such interventions is often to reduce inequalities between individuals by increasing the standing of those in the below

average range. The extent to which interventions can achieve this goal is overestimated by researchers and others who are unaware of the effects of genetic differences between individuals in producing these differences.

The degree to which individual differences on a trait are determined by genetics is expressed as the heritability coefficient, which measures the extent to which the *variance* in the trait across individuals is due to genetic differences between individuals (e.g. cf. Plomin, Defries et al., 2013; Plomin, et al., 2016). For example, if heritability is .50, then 50% of the variance between individuals on this trait is due to genetic differences. The observed heritability of a typical personality trait (e.g., Conscientiousness) is about .50 (Bouchard, 1994, 1997b; Holden, 1987). The square root of the heritability coefficient is the correlation between observed personality test scores and genes; for a heritability of .50, this square root is .71. This correlation is biased downward by measurement error in the personality scale measuring Conscientiousness (e.g, see Nunnally, 1967). Dividing this correlation by the square root of the reliability coefficient of the personality scale removes this bias. This reliability is, at most, .85 when properly computed (cf. Schmidt, Le, & Illies, 2003). Making this bias correction (.71/.92) yields a correlation of .77. A correlation this large between a trait and underlying genes constrains the ability of any intervention to reduce individual differences (i.e., to reduce inequality among individuals). Many such correlations are larger than this. For example, if observed heritability is .60, the correlation is .84. If observed heritability is .70, the correlation is .91. So it is clear that this is not a trivial constraint.

One example of this problem is the *No Child Left Behind Act* passed during the George W. Bush administration. This law had the goal of reducing inequality in academic achievement among

school children by raising achievement levels among students scoring below average for their grade level. Academic achievement, as measured by standardized achievement tests, is a strong function of GMA, with correlations between the two ranging from .77 to .94 (Kaufman, Reynolds, Liu, Kaufman, and McGrew, 2012). Deary, Strand, Smith, & Fernandes (2007) examined the relation between GMA at age 11 and overall academic achievement at age 16 in 25 national academic subject examinations for over 70,000 English children. The correlation, corrected for measurement error, was found to be .81. And GMA is itself highly heritable, with heritability coefficients ranging up to .80 in studies of identical twins separated at birth and reared apart (Bouchard, 2004, 1997a, 1997b). Further, educational achievement itself is highly heritable (.88; Bouchard, 1997b). Academic motivation is often discussed as a cause of academic achievement. But a study based on over 13,000 pairs of identical and fraternal twins from six countries found that the *motivation to achieve academically* also has a large genetic component (Kovas, Garon-Carrier, Boivin, Petrill et al., 2015). In addition, neuroscience research has identified molecular gene sequences that are correlated with educational achievement and attainment (cf. Belsky, Moffitt, Corcoran, et al., 2016; Davies, Marioni, Leiwald, Hill, Hagenaars, Harris et al., 2016). The implication of this research is that it is going to be difficult, if not impossible, for educational interventions to decrease individual differences in academic achievement.

I should be clear that while heritability places constraints on the reductions in individual differences that are possible, it does not necessarily directly limit the extent to which overall means can increase or decrease. It is theoretically possible for the overall mean for a trait or behavior to increase or decrease with no reduction in individual differences on the trait or

behavior. Heritability does not address the overall mean of a trait; it refers only to the differences between individuals in levels of a trait. Interventions of all kinds have failed to lastingly reduce these differences (Scar, 1989).

Another social consequence is the undermining of the credibility of research, especially among informed individuals who are aware that correlation is not causation and/or are aware of some of the facts revealed by behavior genetics. This threat to credibility could generalize to all psychological research studies and is in addition to the threat caused by the questionable statistical data analysis methods that have been discussed at length recently in the literature (e.g, Ioannidis, 2005; Simmons, Nelson, & Simonsohn, 2011).

Failure to Acknowledge Well Established Findings on Specific Abilities and General Mental Ability (GMA)

The second area of problem research focuses on individual aptitudes or specific abilities, such as verbal ability, spatial ability, working memory, etc. Each such aptitude or specific ability is an indicator variable for general mental ability (GMA; Carroll, 1993; Cattell, 1971; Gustafsson, 1984; Jensen, 1998; Schmidt, 2011). There are many such GMA indicator variables with varying levels of quality (i.e., information value about GMA). [Even general knowledge can be a strong indicator variable if comprehensively measured across many knowledge domains (Roznowski, 1997; Schmidt, 2011).] The larger the number of such indicator variables contained in a measure of GMA, the more complete and construct valid is the measure of GMA. After proper statistical control for GMA, the component of each of these abilities that remains adds

essentially nothing to the prediction of real life performances (on the job, in school, in vocational training programs, etc.; Brown, Schmidt & Le, 2006; Ree & Earles, 1991, 1992; Ree, Earles, & Teachout, 1994; Schmidt, Ones, & Hunter, 1992; Thorndike, 1985, 1986). It is only the GMA component in the specific ability that creates its ability to predict life outcomes. That GMA component in the specific aptitude exists because GMA is one cause of the specific aptitude.

Yet many published studies argue that a particular specific ability uniquely predicts life outcomes, without any reference to GMA. For example, some studies conclude that spatial aptitude predicts performance in STEM (science, technology, engineering and math) fields. Some such studies make no mention of a control for GMA. And some present a partial and incomplete "control" for GMA. For example, a short test of vocabulary (e.g, 10 items) might be used to represent GMA. There are two problems with this practice. First, this is only one of the many indicator variables for GMA, so this measure is a construct-deficient measure of GMA. A construct valid measure of GMA requires use of multiple indicator variables (e.g., cf. Carroll, 1993; Gustafson, 1984; Jensen, 1998; Schmidt, 2011). Second, such studies do not control for the biasing effects of measurement error, effects that have long been known (Guilford, 1954; Lord & Novick, 1968; Magnusson, 1966; Nunnally, 1967) and which have been reiterated and re-emphasized in the literature more recently (Brown et al., 2006; Schmidt et al., 2003; Schmidt & Hunter, 1996, 1999, 2015). All measures contain measurement error (Schmidt, 2010), so even a construct valid measure of GMA is not free of measurement error. Without a correction for the biasing effects of measurement error, the control for GMA is only partial. That is, only part of the effect of GMA is removed from the specific aptitude measures in the simultaneous

regression equation; the rest of the GMA effect remains, causing a false indication that the specific aptitude makes a contribution to prediction beyond the effect of GMA (cf. Brown, et al., 2006). We have found no studies of this sort that have employed an adequate control for the effect of GMA on the specific aptitude. And none of them discuss the biasing effects of measurement error.

One example of this occurs in the Study of Mathematically Precocious Youth project (SMPY; Clynes, 2016; Lubinski & Benbow, 2006; Lubinski, Benbow, & Kell, 2014; Lubinski, Benbow, Webb, & Bleske-Recheck, 2016; Matther, Harrison, Lubinski, Putallaz, & Benbow, 2016), which concludes that spatial ability adds incremental prediction over and above GMA. However, the GMA measure used contains only two indicator variables for GMA (verbal and quantitative). A construct-valid GMA measure requires more than two indicators variables. The small observed increment in prediction from the addition of the spatial measure (Clynes, 2016) is what one would expect from adding a third GMA indicator variable, thereby increasing the construct completeness of the overall GMA measure. Finally, in these analyses, there is no correction for the biasing effects of measurement error (Brown et al., 2006; Schmidt et al., 2003; Schmidt & Hunter, 2015).

The investment theory of intelligence (c.f., Cattell, 1971; Schmidt, 2011; 2014) explains why multiple indicators of GMA are required to obtain a complete, construct-valid measure of GMA. Individual interests and individual differences in interests develop relatively early in life and are quite stable over time (Holland, Fritzsche, & Powell, 1994; Kuder & Zytowski, 1988; Low, Yoon, Roberts, & Rounds, 2005; Scarr, 1989; Schmidt, 2011; 2014) and also are substantially heritable (heritability is about .45; Bouchard, 1997b.) An individual's interests determine which specific

abilities or aptitudes the individual invests his or her GMA in developing. For example, someone with strong literary interests will read extensively and invest much of his or her GMA in the development of verbal ability—vocabulary, reading comprehension, literature, writing skills. Someone with strong practical or technical interests will invest more GMA in the development of spatial and mechanical ability. Someone with strong scientific interests will invest more GMA in acquisition of general scientific knowledge (a good indicator of GMA; cf. Schmidt, 2011) and often mathematical and spatial ability. Because different individuals invest their GMA in the development of different specific abilities or aptitudes, it is important to sample multiple indicators of GMA (corresponding to the interests that vary across individuals) in order to achieve a GMA measure that truly represents GMA; that is, a construct-valid measure of GMA (Cattell, 1971; Gustafsson, 1984; Jensen, 1984; 1998; Schmidt 2011).

GMA is not a speculative or merely hypothetical construct. It is arguably the most well established scientific construct in psychology and is comparable in standing to constructs in the physical sciences (Bouchard, 2014; Deary, 2002; Jensen, 1984). GMA is strongly genetically influenced, with a high heritability of about .75 (Bouchard, 2004, 1997a, 1997b; Bouchard & McGue, 1981; McGue & Bouchard, 1998). [Interestingly, the evidence indicates that the more culture-dependent an ability is, the higher its heritability is (Kan, Wicherts, Dolan, & van der Mass, 2013), which is counter to decades of theorizing about culture free and culture fair tests.] At the molecular genetic level, DNA markers have been identified that account for most of the influence of genes on GMA (Plomin, Haworth, Meaburn et al., 2013). Also, there is apparently a physical basis for GMA in the brain: neuroscience research has found important differences in brain structure and function between people who score high and low on GMA measures

(Basten, Hilger, & Fiebach, 2015; Duncan, Seitz, Kolodny et al., 2000; Haier, 2016; Haier & Jung, 2004; Jung & Haier, 2007). GMA measures taken on high school seniors and others aged 15 to 22 predict job level attained 14 years later (Wilk, Desmarais, & Sackett, 1995; Wilk & Sackett, 1996). And in addition to predicting specific occupational, job, and school performances (Schmidt & Hunter, 2004), GMA also predicts the ability of people to perform non-job tasks in everyday life (e.g., reading bus schedules and filling out required forms (Gottfredson, 1997a, 2002). In predicting job performance and performance in occupational training programs, GMA has been shown to be much more important than personality traits (Schmidt & Hunter, 2004; Schmidt, Shaffer, & Oh, 2008). These and other research findings (cf. Gottfredson, 1997b) support the reality of the GMA construct and its central role in human psychology. This support is why published studies that simply ignore GMA while focusing on specific aptitudes and abilities (i.e., individual GMA indicator variables) are lacking in intellectual honesty.

Another example of a problem study is Grzywacz, Segel-Karpas, and Lachman (2016). This study examined the relationship between occupational complexity and three specific aptitudes (indicators of GMA): episodic memory, executive functioning, and memory. They found that, in general, people in more complex occupations had higher scores on these three GMA indicators. There is no indication in this article that people who wind up in more complex, mentally demanding occupations tend to be higher in GMA (Schmidt & Hunter, 2004; Wilk, Desmarais, & Sackett, 1995; Wilk & Sackett, 1996) and that their findings merely reflect the effects of GMA differences on their three indicator variables for GMA.

A third example is the study by Ka-Kellams and Lerner (2016). The dependent variable in this study was empathic accuracy in judging others. They referred to their two independent

variables as “thinking styles”. They state that some people employ an intuitive, emotion-based style in judging others and other people employ a rational style based on reasoning. They found that people who employed a rational, reasoning approach were more accurate in their perceptions of others. There is no mention of the fact that reasoning is an indicator variable for GMA and that it is likely the GMA component captured in reasoning that is responsible for their finding. In fact, the article contains no mention at all of GMA.

One indicator variable for GMA (i.e., one specific aptitude) is working memory, and it is a quality indicator because its correlation with GMA is quite high (e.g., see Colom, Rebollo, Palacios, Juan-Espinosa, & Kyllonen, 2004). Working memory is measured as the ability to hold multiple items of information or objects in mind and to call them up to attention as needed in problem solving. Working memory has often been studied without any reference to its relationship to GMA, and studies have shown that working memory scores can be improved by specific training exercises, as is the case with other indicator variables for GMA (e.g., solving math or spatial problems). Such improvements in specific aptitudes have been found to not generalize or transfer to other aptitudes or to GMA itself. Some studies (e.g., Jaeggi, Buschkuchl, Jonidas, & Perrig, 2008) have claimed that improvements in working memory scores produced by training lead to improvement in GMA. Chooi and Thompson (2012) conducted a careful, well designed study attempting to replicate these claims and found no evidence for such an effect. Recently, Melby-Lervag, Bedick and Hulme (2016) conducted a comprehensive meta-analysis of this literature based on 147 study outcomes. This meta-analysis found that working memory training does not improve performance on measures of GMA or on any other measures of “far transfer” (transfer to other abilities or mental tasks not

identical to working memory tasks). If those involved in research on the working memory aptitude had initially recognized that working memory measures are simply one additional indicator of GMA a great deal of unnecessary research effort could have been avoided.

Studies of the sort we describe here do not present or discuss the well-established research findings on the hierarchical organization of mental abilities (with GMA at the top as the highest and general factor; Carroll, 1993; Gustafsson, 1984; Jensen, 1998). Nor do they discuss the well-established finding that specific GMA indicators contribute essentially nothing to prediction beyond GMA.

Studies like this are lacking in intellectually honesty because both the research summarized in this paper on GMA and mental abilities is clearly available in the literature and is ignored, albeit typically unknowingly; and because the biasing effects of measurement error and the methods for correcting for these biases have long been well known (Guildford, 1954; Lord & Novick, 1969; Magnusson, 1966; Nunnally, 1967) and have been reiterated and emphasized in the literature more recently (Brown et al., 2006; Schmidt & Hunter, 1996; 1999, 2015).

Making the problems in this area clear to the field of psychology is important because there are a large number of such deficient studies in the psychological research literature; these studies are not only misleading in their conclusions, they reduce the general credibility of research on mental abilities.

Discussion

The purpose of this paper is to draw attention to an important problem in the research literature beyond the issue of questionable statistical data analysis methods in research. That is

the problem of the apparent unintentional intellectual dishonesty in many published research articles. I have focused on this problem in two areas. The first area consists of studies purporting to demonstrate the effects of people's experiences on their life outcomes. Many studies drawing causal conclusions about these experiences fail to even mention the probable causal role of genetic inheritance in producing these effects, despite the overwhelming evidence for this connection from behavior genetics research. The second area consists of studies of specific aptitudes (specific abilities) such as verbal, quantitative, and spatial, or reasoning. Many such studies fail to acknowledge or even mention that such aptitudes are indicator variables for general mental ability (GMA) and that after proper control for GMA the residuals in these aptitudes make no contribution to prediction of real world academic, occupational, or job performance. It is only the GMA component in such aptitudes that produces the ability to predict. The issue of the credibility of research conclusions is prominent today (Ioannidis, 2005). In both the areas examined in this paper, these deficiencies create serious and unnecessary credibility problems, and the doubts they inspire about research credibility could unfortunately be generalized unfairly to other research areas in which these problems do not exist. So it is important that this problem be corrected.

However, I do not wish to imply that all studies have these problems. In our search of the literature, we found studies that were exemplary, and for the sake of illustration we describe two here. The study by Dineacu, Turkheimer, Beam, Horn, Duncan & Emery (2016) examined the hypothesis that when people marry, their level of alcohol consumption decreases. They acknowledged that people who marry might be genetically different from those who don't, and they controlled for genetic effects using a sample of 1703 MZ and 722 DZ twins. Their results

showed the even after controlling for genetic effects, people drank less after marriage than they had before marriage. Another exemplary study is Gotlib, LeMoult, Collich, Foland-Ross et al. (2015). This study found that apparently healthy women whose mothers suffered from depression had shorter telomeres and greater cortisol reactivity to stress than did women whose mothers had never been depressed. They acknowledged that this effect could be genetic or environmental or both, and they called for research to determine the relative contributions of genetic and environmental causes. There is another particularly important example. For over 25 years Terrie Moffitt and Avshalom Caspi have conducted longitudinal research studies that have taken into account both genetic and environmental contributions to human behaviors. And in 2016 they received the Distinguished Scientific Contributions Award for this work from the American Psychological Association (Author, 2016). These and other examples in the literature illustrate that it is quite possible to conduct intellectually honest research of the type that I call for in this paper.

Conclusion

It is my hope that the facts this paper has presented will lead to recognition in the literature of the role of behavior genetics findings in studies interpreting correlations between experiences and later life outcomes; and to recognition of the central role of GMA in studies examining specific aptitudes and abilities. At present the literature in these two areas contains many studies that are not scientifically credible. These changes are critical to establishing the credibility of research conclusions in these areas and may help to deter credibility losses across other areas of research.

REFERENCES

Author (2010). January 13, 1998: Georgia Governor Zell Miller proposes writing "The Mozart effect" into law. History.com website. http:www.history.com/this day-in-history/Georgia-governor-zell-miller-proposes-writing-the-mozart-effect-into-law

Author (2016). Award for Distinguished Scientific Contributions: Terrie E. Moffitt and Avshalom Caspi. *American Psychologist, 71,* 658 – 662.

Bank, L., Burraston, B., & Snyder, J. (2004). Sibling conflict and ineffective parenting as predictors of adolescent boys' antisocial behavior and peer difficulties: Additive and interactive effects. *Journal of Research on Adolescence, 14,* 99 – 125.

Basten, U., Hilger, K., & Fiebach, C. J. (2015). Where smart brains are different: A quantitative meta-analysis of functional and structural brain imaging studies on intelligence. *Intelligence, 51,* 10 – 27. https://doi.org/10.1016/j.intel.2015.04.009

Baumrind, D. (1993). The average expectable environment is not good enough: A response to Scarr. *Child Development, 64,* 1299 – 1317.

Baumrind, D. (2009). Necessary distinctions. *Psychological Inquiry, 8,* 176 – 229. http://dx.doi.org/10.1207/s15327965pli0803_2

Belsky, D. W., Moffitt, T. E., Corcoran, D. L. et al. (2016). The genetics of success: How single nucleotide polymorphisms associated with educational attainment relate to life course development. *Psychological Science Online First.* doi: 10.1177/0956797616643070.

Bouchard, T. J., Jr. (2014). Genes, evolution, and intelligence. *Behavior Genetics, 44,* 549 – 577. doi: 10.1007/s10519-014-9646-x

Bouchard, T. J., Jr. (2004). Genetic influences on human psychological traits: A survey. *Current Directions in Psychological Science, 13,* 148 – 151.

Bouchard, T. J., Jr. (1997a). IQ similarity in twins reared apart: Findings and responses to critics. In R. J. Sternberg and E. L. Grigorenko (Eds.), *Intelligence: Heredity and environment.* NY: Cambridge, pp. 126 – 160.

Bouchard, T. J., Jr. (1997b). Genetic influences on mental abilities, personality, vocational interests, and work attitudes. *International Review of Industrial and Organizational Psychology, 12,* 373 – 395.

Bouchard, T. J., Jr. (1994). Genes, environment, and personality. *Science, 264,* 1700 – 1701.

Bouchard, T. J., Jr. & McGue, M. (1981). Familial studies of intelligence: A review. *Science, 212,* 1055 – 1059.

Brown, B. B., Mounts, N. S., Lamborn, S. D., & Steinberg, L. (1993). Parenting practices and peer group affiliation in adolescence. *Child Development, 64,* 467 – 347.

Brown, K., Schmidt, F. L., & Le, H. (2006). Specific aptitude theory revisited: Is there incremental validity for training performance? *International Journal of Selection and Assessment, 14,* 87 – 100.

Campbell, D. (2001). *The Mozart Effect: Tapping the power of music to heal the body, strengthen the mind, and unlock the creative spirit.* NY: Harper.

Carretta, T. R., & Ree, M. J. (1995). Near identity of cognitive structure in sex and ethnic groups. *Personality and Individual Differences, 19,* 149 – 155.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies.* Cambridge University Press: Cambridge, UK.

Carroll, J. B. (1997). Psychometrics, intelligence, and public perception. *Intelligence, 24,* 25 – 52.

Cattell, R. B. (1971). *The investment theory of intelligence.* NY: Wiley.

Chen, E., & Miller, G. E. (2012). "Shift-and-Persist" strategies: Why low socioeconomic status isn't always bad for health. *Perspectives on Psychological Science, 7,* 135 – 158.

Chen, E., Turiano, N. A., Mroczek, D. K. et al. (2016). Association of reports of childhood abuse and all-cause mortality rates in women. *JAMA Psychiatry, 73,* 920 – 927. doi:10.1001/jamapsychiatry.2016.1787

Chooi, W-T, & Thompson, L. A. (2012). Working memory training does not improve intelligence in healthy young adults. *Intelligence, 40,* 531 – 542. http://dx.doi.org/10.1016/jintel2012.07.004

Clynes, T. (2016). How to raise a genius. *Nature, 537,* 152 – 155.

Colarelli, S. M., & Arvey, R. D. (2015). *The biological foundations of organizational behavior.* Chicago: University of Chicago Press.

Collins, W. A., Maccoby, E. E., Steinberg, L., Hetherington, E. M., & Bornstein, M. H. (2000). Contemporary research on parenting: The case for nature and nurture. *American Psychologist, 55,* 218 – 232.

Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., & Kyllonen, P. C. (2004). Working memory is (almost) perfectly predicted by *g*. *Intelligence, 32,* 277 – 296. Doi:10.1016/j.intell.2003.12.002.

Daly, M., Egan, M., Quigley, J., Delaney, L., & Baumeister, R. F. (2016). Childhood self-control predicts smoking through life: Evidence from 21,000 cohort study participants. *Health Psychology, 35,* 1254 – 1263. http//dx.doi.org/10.1037/hea0000393

Darling, N., & Steinberg, L. ((1992). Parenting style as context: An integrative model. *Psychological Bulletin, 113,* 487 – 496.

Davies, G., Marioni, R. E., Liewald, D. C., Hill, W. D., Hagenaars, S. P., Harris, S. E., et al. (2016). Genome-wide association study of cognitive functions and education attainment in UK Biobank (N = 112151). *Molecular Psychiatry, 21,* 758 – 767. https://doi.org/10.1038/mp.2016.45

Davis, A. N., Carlo, G., Schwartz, S. J., Unger, J. B. et al. (2016). The longitudinal associations between discrimination, depressive symptoms, and prosocial behaviors in U.S. Latino/ recent immigrant adolescents. *Journal of Youth and Adolescence, 45,* 457 – 470.

Deary, I. J. (2002). The general factor of intelligence. In Sternberg, R. J., & Grigorenko, E. L. (Eds.), *Intelligence: Measurement and Theory.* Mahwah, NJ: Erlbaum, pp. 151 – 182.

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence, 35,* 13 – 21.

DeBaryshe, B. D., Patterson, G. R., & Capaldi, D. M. (1993). A performance model for academic achievement in early adolescent boys. *Developmental Psychology, 29,* 795 – 804.

Dinescu, D., Turkheimer, E., Beam, C. R., Horn, E. E., Duncan, G., & Emery, R. E. (2016). Is marriage a buzzkill? A twin study of marital status and alcohol consumption. *Journal of Family Psychology, 30,* 698 – 707. http//dx.doi.org/10.1037/fam0000221 (example of good study—controlled for genetics)

Duncan, J., Seitz, R. J., Kolodny, J., Bor, D., Herzog, H. & Ahamed, A. et al. (2000). A neural basis for general intelligence. *Science, 289,* 457 – 460.

Durbin, D. L., Darling, N., Steinberg, L., & Brown, B. B. (1992). Parenting style and peer group membership among European-American adolescents. *Journal of Research on Adolescence, 3,* 87 – 100.

Friedman-Krauss, A. H. & Raver, C. C. 2015). Does school mobility place elementary school children at risk for lower math achievement? The mediating role of cognitive dysregulation. *Developmental Psychology, 61, 1725 – 1739.* http//ds.doi.org/10.1037/a0038795

Gotlib, I. H., Lemoult, J., Colich, N. L. et al. (2015). Telomere length and cortisol reactivity in children of depressed mothers. *Molecular Psychiatry, 20,* 615 – 620. (example of good study)

Gottfredson, L. S. (1997a). Why *g* matters: the complexity of everyday life. *Intelligence, 24,* 79 – 132.

Gottfredson, L. S. (1997b). Mainstream science on intelligence: An editorial with 52 signatories, history and bibliography. *Intelligence, 24,* 13 – 23. https://doi.org/1016/S0160-2896(97)90011-8

Gottfredson, L. S. (2002). Where and why *g* matters: Not a mystery. *Human Performance, 15,* 25 – 46.

Grzywacz, J. G., Segel-Karpas, D., & Lachman, M. E. (2016). Workplace exposures and cognitive function during adulthood: Evidence from National Survey of Midlife Development and the O*Net. *Journal of Occupational and Environmental Medicine, 58,* 535 – 541. Doi:10.1097/JOM.0000000000000727

Guilford, J. P. (1954). *Psychometric Methods.* NY: McGraw-Hill.

Gustafsson, J.-E. (1994). A unifying model of the structure of intellectual abilities. *Intelligence, 8,* 179 – 203.

Haier, R. J. (2016). *The neuroscience of intelligence.* NY: Cambridge University Press.

Haier, R. J., Jung, R. E., et al. (2004). Structural brain variation and general intelligence. *NeuroImage, 23,* 425 – 433.

Hair, N. L., Hanson, J. I., Wolfe, B. L., & Pollak, S. D. (2015). Association of child poverty, brain development, and academic achievement. *JAMA Pediatrics*, *169,* 822 – 829. Doi.10.1001/jamapediatrics.2015.1475

Hetland, L. (2000). Listening to music enhances spatial-temporal reasoning: Evidence for the Mozart Effect. *Journal of Aesthetic Education, 34,* 105 – 148.

Holden, C. (1980). Reunited twins: More than the faces are similar. *Science Magazine, 80,* 55 – 59.

Holden, C. (1987). The genetics of personality. *Science, 237,* 598 – 601.

Holland, J. L., Fritzsche, B., & Powell, A. (1994). *Self-directed search: Technical manual.* Odessa, FL: Psychological Assessment Resources.

Houston, A. C., McLoyd, V. C., & Garcia-Coll, C. (1994). Children and poverty: Issues in contemporary research. *Child Development, 65,* 275 – 282.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *Chance, 18,* 40 – 47. doi: 10.1080/09332480.10722754

Jaeggi, S. M., Buschkuehl, M., Jonidas, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *PANAS: Proceedings of the National Academy of Sciences of the United States of America, 105,* 6829 – 6833. doi: 10.1073/pnas.0801268105.

Jensen, A. R. (1984). Test validity: *g* vs. the specificity doctrine. *Journal of social and biological structures, 7,* 93 – 118.

Jensen, A. R. (1998). *The g factor: The science of mental ability.* Westport, CT: Praeger.

Judge, T. A., Klinger, R. L., & Simon, L. S. (2010). Time is on my side: Time, general mental ability, human capital, and extrinsic career success. *Journal of Applied Psychology, 95,* 92 – 107.

Jung, R. E., & Haier, R. J. (2007). The parieto-frontal integration theory (P-Fit) theory of intelligence: Converging neuroimaging evidence. *Behavioral and Brain Sciences, 30,* 135 – 154.

Kan, K-J., Wicherts, J. M., Dolan, C. V., & van der Maas, H. L. J. (2013). On the nature and nurture of intelligence and specific cognitive abilities: The more heritable, the more culture dependent. *Psychological Science, 24,* 2420 – 2428. doi: 10.1177/0956796113493292

Kaufman, S. B., Reynolds, C. R., Liu, X., Kaufman, A. S., & McGrew, K. S. (2012). Are cognitive *g* and academic Achievement *g* one and the same? An exploration on the Woodcock-Johnson and Kaufman Tests. *Intelligence, 40,* 123 – 138.

Kaufman, J., & Zigler, E. (1987). Do abused children become abusive parents? *American Journal of Orthopsychiatry, 57,* 186 – 192. doi: 10.111/jj.1939-0025.1987.tb03528x

Kaufman, J., & Zigler, E. (1989). The intergenerational transmission of child abuse. In D. Cichetti & V. Carlson (Eds.), *Child maltreatment: Theory and research on the causes and consequences of child abuse and neglect.* UK: Cambridge University Press, pp. 129 – 152.

Kovas, Y., Garon-Carrier, G., Boivin, M., Petrill, S. A. et al. (2015). Why children differ in motivation to learn: Insights from over 13,000 twins from 6 countries. *Personality and Individual Differences, 80,* 51 – 63.

Lee, J. J., & McGue, M. (2016). Why behavior genetics matters: Comment on Plomin et al. (2016). *Perspectives on Psychological Science, 11,* 29 – 30.

Lippold, M. A., Davis, K. D., McHale, S. M., Buxton, S. M., Orfeu, M., & Almeida, D. M. (2016). Daily stressor reactivity during adolescence: The buffering role of parental warmth. *Health Psychology, 35,* 1027 – 1035.

Lord, C., Cook, E. H., Leventhal, B. L., & Amaral, D. C. (2000). Autism spectrum disorders. *Neuron, 28,* 355 – 363. doi: org/10.1016/SO896-6273(00)00115-X

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores.* London, UK: Addison-Wesley.

Low, K. S. D., Yoon, M., Roberts, B. W., & Rounds, J. (2005). The stability of vocational interests from early adolescence to middle adulthood: A quantitative review of longitudinal studies. *Psychological Bulletin, 131,* 713-737.

Lubinski, D., & Benbow, C. P. (2006). Study of mathematically precocious youth after 35 years: Uncovering the antecedents for the development of math-science expertise. *Perspectives on Psychological Science, 1,* 316 - 345.

Lubinski, D., Benbow, C. P., & Kell, H. J. (2014). Life paths and accomplishments of mathematically precocious males and females four decades later. *Psychological Science, 25,* 2217 – 2232.

Lubinski, D., Benbow, C. P., Webb, R. M., & Bleske-Rechek, A. (2016). Tracking exceptional human capital over two decades. *Psychological Science, 17,* 194 – 199.

Mackenzie, D. (1999). Requiem for the Mozart Effect? AAAS *Science,* July Issue.

MacKinnon, C., Starnes, R., Volling, B., & Johnson, S. (1997). Perceptions of parenting as predictors of boy's sibling and peer relationships. *Developmental Psychology, 33,* 1024 – 1031. doi.org/10.1037/0012-1649.33.6.1024

Magnusson, D. (1966). *Test Theory.* NY: Addison-Wesley.

Makel, M. C., Harrison, J. K., Lubinski, D., Putallaz, M., & Benbow, C. P. (2016). When lightning strikes twice: Profoundly gifted, profoundly accomplished. *Psychological Science, 27,* 1 – 15. doi: 10.1177/095679761664735

Ma-Kellams, C., & Lerner, J. (2016). Trust your gut or think carefully? Examining whether an intuitive versus a systematic, mode of thought produces greater empathic accuracy. *Journal of Personality and Social Psychology, 111,* 674 – 685. http//dx.doi.org/10.1037.pspi0000063

Margolis, A. E., Herbstman, J. B., Davis, K. S., Thomas, V. K. et al. (2016). Longitudinal effects of prenatal exposure to air pollutants on self-regulatory capacities and social competence. *Journal of Child Psychology and Psychiatry, 57,* 851 – 860. doi: 10.1111/jcpp.12548

McGue, M., & Bouchard, T. J., Jr. (1998). Genetic and environmental influences on human behavioral differences. *Annual Review of Neuroscience, 21,* 1 – 24.

McLearn, G. E. (1997). Substantial genetic influences on intelligence in twins 80 years old or older. *Science, 276,* 1560 – 1563.

Melby-Lervag, M., Redick, T. S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of "far transfer": Evidence from a meta-analytic review. *Psychological Science, 27,* 512 – 534.

Nunnally, J. C. (1967). *Psychometric theory.* NY: McGraw-Hill.

Petrill, S. A., Plomin, R., Berg, S., Johansson, B., Pedersen, N. L., Ahern, G., & McClearn, G. E. (1998). The genetic and environmental relationship between general and specific abilities in twins age 80 and older. *Psychological Science, 9,* 183 – 189.

Plomin, R., & Bergman, C. S. (1991). Nature and nurture: Genetic influences on "environmental" measures. *Behavioral and Brain Sciences, 14,* 373 – 427.

Plomin, R., Haworth, C. M. A., Meaburn, E. L., Price, T. S., Wellcome Trust Case Control Consortium 2, & Davis, O. S. P. (2013). Common DNA markers can account for more than half of genetic influence cognitive abilities. *Psychological Science, 24,* 562 – 568.

Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2013). *Behavioral genetics* (6th ed.). New York, NY: Worth.

Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2016). Top 10 replicated findings from behavior genetics. *Perspectives on Psychological Science, 11,* 3 – 23.

Plomin, R., Owen, M. J., & McGuffin, P. (1994). The genetic basis for complex human behaviors. *Science, 264,* 1733 – 1739.

Porcerelli, J. H., Huth-Bocks, A., Huprich, S. K., & Richardson, L. (2016). Defense mechanism of pregnant mothers predict attachment security, social-emotional competence, and behavior problems in their toddlers. *Psychiatry, 173,* 138 – 146.

Ree, M. J., & Earles, J. A. (1992). Intelligence is the best predictor or job performance. *Current Directions in Psychological Science, 1,* 86 – 89.

Ree, M. J., & Earles, J. A. (1991). Predicting training success: Not much more than *g. Personnel Psychology, 44, 321 – 332.*

Ree, M. J., Earles, J. A., Teachout, M. S. (1994). Predicting job performance: Not much more than *g*. *Journal of Applied Psychology, 79,* 518 – 524.

Rowe, D. C. (1994). *The limits of family influence: Genes, experiences, and behavior.* NY: Guilford Press.

Rowe, D. C. (2001). The nurture assumption persists. *American Psychologist, 56,* 168 – 174.

Roznowski, M. (1987). Use of tests manifesting sex differences as measures in intelligence: Implications for measurement bias. *Journal of Applied Psychology, 72,* 480 – 483.

Salgado, J. F., Anderson, N., Moscoso, S., Betua, C., & de Fruyt, F. (2003). International validity generalization of GMA and cognitive abilities: A European Community meta-analysis. *Personnel Psychology, 56,* 573 – 605.

Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C., de Fruyt, F., & Roland, J. P. (2003). A meta-analytic study of general mental ability for different occupations in the European community. *Journal of Applied Psychology, 88,* 1068 – 1081.

Scar, S. (1996). How people make their own environments: Implications for parents and policy makers. *Psychology, Public Policy, and Law, 2,* 204 – 228.

Scar, S. (1989). Protecting general intelligence. In R. L. Linn (Ed.), *Intelligence: Measurement theory and public policy.* University of Illinois Press. Pp. 74 – 118.

Scar, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype to environment effects. *Child Development, 54,* 424 – 435.

Schmidt, F. L. (2002). The role of general cognitive ability in job performance: Why there cannot be a debate. *Human Performance, 15,* 187 – 210.

Schmidt, F. L. (2010). Detecting and correcting the lies that that data tell. *Perspectives on Psychological Science, 5,* 233 – 242.

Schmidt, F. L. (2011). A theory of sex differences in technical aptitude and some supporting evidence. *Perspectives on Psychological Science, 6,* 650 – 673.

Schmidt, F. L. (2014). A general theoretical integrative model of individual differences in interests, abilities, and academic and occupational achievement: A commentary on four recent articles. *Perspectives on Psychological Science, 9,* 211 – 218.

Schmidt, F. L., & Hunter, J. E. (1998). The validity of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124,* 262 – 274.

Schmidt, F. L., & Hunter, J. E. (1999). Theory testing and measurement error. *Intelligence, 27,* 183 – 198.

Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods, 1,* 199 – 223.

Schmidt, F. L., & Hunter, J. E. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology, 86,* 162 – 173.

Schmidt, F. L., Hunter, J. E., & Caplan, J. R. (1981). Validity generalization results for two job groups in the petroleum industry. *Journal of Applied Psychology, 66,* 261 – 273.

Schmidt, F. L., Le, H., & Ilies, R. (2003). Beyond Alpha: An empirical examination of the effects of different sources of measurement error on reliability estimates for measures of individual differences constructs. *Psychological Methods, 8,* 206 – 234.

Schmidt, F. L., & Oh, I-S. (2016). The crisis of confidence in research findings in psychology: Is lack of replication the real problem? Or is it something else? *Archives of Scientific Psychology, 4,* 32 – 37.

Schmidt, F. L., Schaffer, J. R., & Oh, I-S. (2008). Increased accuracy for range restriction corrections: Implications for the role of personality and general mental ability in job and training performance. *Personnel Psychology, 61,* 827 – 868.

Schmidt, F. L., Oh, I-S., & Shaffer, J. R. Update of The validity of selection methods in personnel psychology: Practical and theoretical implications of 100 years of research findings. Paper under review.

Schmidt, F. L., Ones, D., & Hunter, J. E. (1992). Personnel Selection. *Annual Review of Psychology, 43,* 627 – 670.

Sigman, M., & Capps, L. (1997). *Children with autism: A developmental perspective.* Cambridge: Harvard University Press.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22,* 1359 – 1366. http//dx.doi.org/10.1177/0956797611417632

Steele, K. M., Bass, K. E., & Crook, M. D. (1999). The mystery of the Mozart Effect: Failure to replicate. *Psychological Science, 10,* 366 – 369.

Stoltenberg, S. F., & Burmeister, M. (2000). Recent progress in psychiatric genetics. *Human Molecular Genetics, 9,* 927 – 935. doi: 10.1093/hmg/9.6.927

Tang, S., Davis-Kean, P. E., & Sexton, H. R. (2014). Adolescent pregnancy's intergenerational effects: Does an adolescent mother's education have consequences of her children's achievement? *Journal of Research on Adolescence, 26,* 180 – 193. doi: 10.1111/jora.12182

Taylor, C., Manganello, J. A., Lee, S. J., & Rice, J. C. (2010). Mothers' spanking of 3-year-old children and subsequent risk of childrens' aggressive behavior. *Pediatrics, 125,* 127 – 136.

Thorndike, R. L. (1985). The central role of general ability in prediction. *Multivariate Behavioral Research, 20,* 241 – 254.

Thorndike, R. L. (1986). The role of general ability in prediction. *Journal of Vocational Behavior, 13,* 332 – 339.

Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. *Current Directions in Psychological Science, 5,* 160 – 164.

Wachs, T. D., & Gruen, G. (1982). *Early experience and human development.* NY: Plenum Press.

Waldinger, R. J., & Schulz, M. S. (2016). The long reach of nurturing family environments: Links with midlife emotion-regulatory styles and late-life security in intimate relationships. *Psychological Science, 27,* 1443- 1450.

Waller, N. G., Kojetin, B. A., Bouchard, T. J., Jr., Lykken, D. T., & Telligen, A. (1990). Genetic and environmental influences on religious interests, attitudes, and values: A study of twins reared apart and together. *Psychological Bulletin, 116,* 138 – 142.

Walaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological Science, 25,* 1314 – 1324. doi: 10.1177/0956797614531023

Zheng, Y., Plomin, R., & von Stumm, S. (2016). Heritability in intraindividual mean and variability of positive and negative affect: Genetic analysis of daily affect ratings over a month. *Psychological Science, 27,* 1611 – 1619. doi: 10.1177/095679761666994