Abstract

In this reply to the comments on our original article, we identify common themes and respond to some specific issues that appear to be at the core of the debate over the evidence for “Generation Me.” In particular, we point out that generational studies have both applied and theoretical consequences and echo the sentiment that developmental considerations are likely to be more important than generational considerations when thinking about the attributes of today’s young people. We also acknowledge that disagreements are inevitable, as this is a controversial issue of study. However, we reiterate our commitment to the interpretation of effect sizes rather than null hypothesis significance testing and emphasize the need for care when psychological
scientists offer broad and often moralistic pronouncements about entire generations of young people.

Keywords: Generation Me, Millennials, cohort effects, self-esteem, adolescents

The goal of our original paper was to evaluate evidence for secular changes in the attitudes and personal characteristics of high-school seniors from 1976 to 2006 in line with the recent characterizations of “Generation Me” (those individuals born in the 1970s, 1980s, and 1990s; see Twenge, 2006, p. 3). In light of the attention that generational changes have received in both academic discussions and the popular press, we were especially interested in quantifying the size of generational differences using intuitive descriptive statistics. The range of issues raised in the comments reflects both the interest in the topic as well as the complexities and controversies involved when addressing questions about generational changes. In this short response to the five comments, we highlight some common themes present in the constructive comments and offer a rejoinder to some of the criticisms of our article. Throughout, we attempt to identify some interesting and important areas for future study. First, we issue a clarification about sampling issues as they apply to our original report.

MEA CULPA: OUR SLOPPY WRITING ABOUT SAMPLING AND REPRESENTATIVENESS

After reviewing the comments, we returned to our original article and found that we had been somewhat sloppy in our writing about issues of sampling and generalizability. Indeed, we should have been explicit throughout that the Monitoring the Future (MTF) surveys that we examined were always completed by high-school seniors, a clearly more restrictive subset of the
population than all American young people. Besides the fact that we should have been more circumscribed when talking about the generalizability of the MTF, we should have also been more thorough when discussing issues of sampling. Strictly speaking, the notion of representativeness applies to sampling strategies rather than samples themselves, and we should not have used such shorthand in the original article. Bachman and O’Malley (two of the primary MTF investigators) in a personal communication with us (dated September 21, 2009) noted that they themselves take pains to describe “the various limitations of our particular survey methods in [official publications from the MTF such as Johnston, O’Malley, Bachman, & Schulenberg, 2009, and earlier volumes].” In hindsight, we should have followed the examples set by Bachman and O’Malley more closely when writing about the MTF.

We hope that our omissions do not overshadow important points about the advantages of the MTF project for tracking generational changes. First, the samples were generated using probability methods and strict sampling methodologies, something that is not true of the majority of the samples included in cross-temporal meta-analyses. Although there are inevitable limitations placed on the generalizability of the findings from any study, the relevant issue is which approach is likely to yield a closer approximation of the size of the true changes in the attributes of young people (high-school seniors or otherwise). Second, it is useful to be clear about the impact of the bias in this context. In their personal communication, Bachman and O’Malley made a useful point about biases that we wish to reiterate: “…because our methods are highly consistent across the years, the unavoidable biases are also consistent across time; thus, trend data should be relatively unbiased.” Given these considerations, we continue to believe that the MTF is a very useful resource for testing hypotheses related to Generation Me.
OVERARCHING THEME #1: GENERATIONAL STUDIES OFTEN HAVE IMPORTANT APPLIED AND THEORETICAL IMPLICATIONS

One subtext of several of the comments is that studies of generational change have potentially important applied implications. Our primary applied concern has followed from a conviction that scientific research should not be used to fuel unfounded and negative stereotypes of young people. Indeed, Arnett (2010, this issue) suggested that many popular characterizations of today’s young people are negative and provided an explanation as to why these stereotypes seem to be widespread. The negative impressions of young people are a fairly ancient phenomenon (see e.g., Roberts, Edmonds, & Grijalva, 2010, this issue), and such considerations have made us skeptical of recent claims regarding the seemingly pervasive negative characterizations of this generation of young people. To be sure, studies pointing out generational differences are quick to attract the attention of the media and others outside of academic social science. This suggests to us that special attention is needed to translate the findings, limitations, and practical implications (if any) for lay audiences clearly and appropriately. It is undoubtedly important to give psychological research “back” to the public; however, it is also important to preserve the complexity of the underlying trends and to convey limitations in the current state of knowledge.

Beyond applied concerns, Terracciano (2010, this issue) noted that generational studies have theoretical implications in terms of the power of broader socio-historical contexts for shaping personal attributes, particularly those personality attributes contained in the Five Factor Model (e.g., McCrae & Costa, 2008). At the heart of this issue is the degree to which personality is shaped by cultural and historical factors—an issue that has implications for how personality traits are conceptualized by the field. To that end, Terracciano summarized the results of an
expansive and extensive program of research on adult personality development that includes cross-sectional, longitudinal, and cross-national research. Our broad skepticism about the existence of strong generational effects for personality attributes is consistent with his summary (see also McCrae & Costa, 2003, p. 80). Terracciano also raised an instructive point about the nature of replication. That is, the issue is whether an effect size more or less replicates across studies and not simply whether the direction of an effect is duplicated. If researchers report a strong cohort effect for Extraversion, for example (say e.g., something approaching a 1 SD increase over 40 years), then it seem reasonable to expect that other empirical studies should find similarly large effects for the same trait. Terracciano, however, provided evidence that challenges the robustness of the earliest cross-temporal meta-analyses with respect to apparently dramatic generational increases in Neuroticism and Extraversion.

**OVERARCHING THEME #2: DEVELOPMENTAL CONSIDERATIONS ARE IMPORTANT WHEN CONSIDERING THE ATTRIBUTES OF YOUNG PEOPLE**

Both Arnett (2010) and Roberts et al. (2010) draw attention to the importance of developmental factors, as opposed to purely generational ones, when considering the attributes of young people. We strongly concur with these sentiments and believe that Arnett (2000, 2004, 2007) has provided a useful framework for thinking about personality and social development during the transition to adulthood. His notion of emerging adulthood embodies the inherently developmental insight that the transition from adolescence to adulthood takes time, especially in industrialized societies. A key insight is that those in their late teens to even their late 20s (Arnett, 2007) have particular developmental tasks and they likely have somewhat different psychological attributes when compared with “true” adults (i.e., individuals with clearly defined roles such as parent, full-time employee, or committed romantic partner).
Arnett also provided evidence for increases in well-being that occur during the period of emerging adulthood. We suspect that this developmental finding is not likely to be strongly modified by cohort, given that a recent analysis of the statistical effects of age, period, and cohort on happiness using the General Social Survey (Yang, 2008) also pointed to age-related increases in happiness. Although one could question the importance of even the strongest effect sizes in this study, a basic message was that “age effects dominate and the period and cohort variations are small” (p. 221).

These results for happiness and well-being fit with developmental trends for other personality attributes—characteristics which seem to follow a generalization known as the maturity principle of personality development (Caspi, Roberts, & Shiner, 2005). The basic idea is that there are normative increases in traits relevant to performing the important adult roles of worker, committed partner, and caregiver. Indeed, researchers focusing on personality development have accumulated replicable evidence for increasing levels of constraint and decreasing levels of hostility and alienation from the late teens through the 20s (e.g., Blonigen, Carlson, Hicks, Krueger, & Iacono, 2008; Donnellan, Conger, & Burzette, 2007; Galambos, Barker, & Krahn, 2006; Roberts, Caspi, & Moffitt, 2001; Roberts, Walton, & Viechtbauer, 2006). Our concern is that the largely consistent findings regarding increasing maturity and well-being during the transition to adulthood proper can be lost in recent discussions of the apparently negative characteristics of today’s young people.¹

In fact, we believe that it is quite likely that people become less self-focused as they assume adult roles and responsibilities. In line with this speculation, Roberts et al. (2010) suggested the important possibility of age-related differences in the attributes assessed by the
Narcissistic Personality Inventory (NPI; Raskin & Terry, 1988). They noted that the age differences they found converged with earlier cross-sectional Internet data collected in 2001 and 2002 by Foster, Campbell, and Twenge (2003). The next step should be to design and carry out longitudinal studies that use the NPI or perhaps psychometrically better and more targeted measures of the multifaceted construct of narcissism (see Brown, Budzek, Tamborski, 2009; Pincus et al., 2009) to evaluate whether cross-sectional age differences map to longitudinal changes over the course of the adult life span.²

OVERARCHING THEME #3: THE INTERPRETATION OF GENERATIONAL TRENDS IS CONTENTIOUS

The overall tone of the ongoing discussions about generation trends can get heated as different groups debate the interpretation of particular findings. This is quite clear from the comments. Indeed, Eckersley (2010, this issue) noted that there are proponents of all positions on the broad issue of generational changes. Likewise, we absolutely agree with Eckersley (2010) when he concluded that our original article has not settled debates about generational trends in psychological attributes. Our goal was to evaluate predictions based on Generation Me type accounts of today’s youth, and we concluded that effect sizes are much less dramatic than portrayed in previous discussions. This perspective seems to be consistent with Eckersley’s acknowledgement that “that not all time-series studies have found increased psychological problems” (p. xx).

We do, however, wish to add some additional data to the trends described by Eckersley (2010). He cites a considerable amount of evidence about trends from countries outside of the United States which widened the discussion beyond our original focus on changes in the United States. Recall that some of the proposed origins of Generation Me trends are thought to be found
in seemingly iatrogenic effects of the self-esteem movement, a movement that had its roots in California sometime in the 1980s. Counter to concerns about the deleterious effects of this movement, we are unaware of consistent evidence suggesting that either youth internalizing or externalizing problems in the United States have dramatically increased in recent years following the self-esteem movement.

Violence and crime may serve as a very useful societal-level manifestation of generational trends as they pertain to “Generation Me” claims and related concerns over the consequences of the self-esteem movement. The reality is that rates of violent crime in the United States have typically fallen since the 1990s. For example, the rate of violent crime per 100,000 inhabitants was 454.5 in 2008 as opposed to 758.2 in 1991 (see http://www.fbi.gov/ucr/cius2008/data/table_01.html). These drops also hold for reports of criminal victimization (http://www.ojp.usdoj.gov/bjs/glance/viort.htm or http://childstats.gov/americaschildren/beh5.asp). Thus, the two major ways of assessing crime and violence in the United States, the Uniform Crime Reports and the National Crime Victimization Survey both point to declines in violence during the peak of the alleged epidemic of narcissism (for a similar perspective on crime trends, see Ferguson, 2008). These trends are particularly relevant given that young people account for the lion’s share of criminal offenses (i.e., the age-crime curve; Hirschi & Gottfredson, 1983) and the fact that Twenge has drawn connections between changes in narcissism and violent acts (e.g., Twenge, 2006, pp. 70–71). In short, even if one decides that the cohort-related changes in “Generation Me” traits are substantial, it does not appear that those changes have translated to societal-level increases in violence and aggression, two of the most pernicious outcomes thought to be associated with egotism and narcissism.
**GROUNDHOG DAY AND RED HERRINGS: RESPONDING TO TWENGE AND CAMPBELL (2010, THIS ISSUE)**

Whereas Twenge and Campbell (2010) invoke the image of Alice from *Alice in Wonderland* in response to our original article, their comments made us feel a bit like Bill Murray in *Groundhog Day*. We have now had several exchanges with Twenge and her colleagues covering many of the issues (Donnellan & Trzesniewski, 2009; Donnellan, Trzesniewski, & Robins, 2009; Trzesniewski & Donnellan, 2009; Trzesniewski, Donnellan, & Robins, 2008) and this back-and-forth seems to generate far more heat than light. In this response, we focus on just a few broad issues and direct interested readers to these other publications for more details (especially Donnellan & Trzesniewski, 2009).

Foremost, we did not use an illogical approach to analyzing the MTF data. We followed the logic of practices outlined by Jacob Cohen (1988) to evaluate the importance of the current findings (see Cohen, 1988, p. 16-17). In particular, we focused on effect sizes as opposed to adopting the ritual of null hypothesis statistical significance testing. The size of the MTF dataset guarantees that all but the tiniest of effects reaches a $p$ value of less than .01 (see Trzesniewski & Donnellan, 2009). Cohen himself noted that it is “ridiculous to try to develop theories about human behavior with $p$ values from Fisherian hypothesis testing and no more than a primitive sense of effect sizes” (1990, p. 1309). After decades of controversy, the limitations of null hypothesis statistical significance testing and potential advantages of an explicit focus on effect sizes should be well known to psychologists (e.g., Ferguson, in press; Kline, 2004; see Fraley & Marks, 2007, for a review). Rather than blind readers with the asterisks denoting statistical significance, we provided several options for gaining a sense of the magnitude of the generational changes in the measures that we examined: means by 5 year intervals, correlations
between year and level of a construct, and regression coefficients (in an unstandardized metric). These are all clearly reported in Table 3 of our original article. We hope that readers share our convictions about the importance of effect sizes as opposed to $p$ values for the practice of scientific psychology (see also American Psychological Association, 2009, p. 34; Kashy, Donnellan, Ackerman, & Russell, 2009, pp. 1138–1139).

Of course, just like the alpha level of statistical significance test is a judgment call, so are judgments about the size of the differences in attributes across generations. Regarding the latter, reasonable minds can disagree; however, we acknowledged that such interpretations involve subjectivity and we spent a considerable amount of time justifying our interpretation of effect sizes in our original article. Our decisions were fairly conventional and they were not made on a whim. We can even note that others like Ferguson (in press) suggest even more stringent rules of thumb for the interpretation of effect sizes.

Beyond the issue of significance testing, we would like to clarify a few things about the MTF dataset, given some of the concerns raised about this resource. Accordingly, we sought the advice of Jerald Bachman and Patrick O’Malley, two of the investigators of the MTF study. They agreed that “it is important to correct some misconceptions about the Monitoring the Future (MTF) surveys in general, and about MTF measures of self-esteem in particular” (personal communication, September 21, 2009). Twenge and Campbell raised the possibility that changes in item ordering may affect the ability of researchers to detect time trends. Bachman and O’Malley acknowledged that:

“Because some of the issues studied over the years have changed, it has been necessary on occasion to make changes in item content and item order. However, we have been quite mindful of the risks of changing the contexts in which items appear, and thus take
considerable pains to minimize such changes, and to check on possible effects when such changes do occur. In the specific case of self-esteem measures, we have added the self-esteem items to additional forms over the course of time in order to permit a wider range of correlational analyses. As it turns out, our own analyses at several times have revealed little effect of item context on the self-esteem measure, particularly when examining relations with other items (including demographic and other measures that appear in all forms). In any case, however, the proper way to measure trends on self-esteem using MTF data would be to limit analyses to Form 5, in which the same set of items (as well as the preceding items, which provide the context) have appeared unchanged from 1977 onward. Moreover, to minimize random fluctuations due to sample size limitations in any one year (particularly for single-form measures), it is far preferable to analyze data across multiple years; ideally, data from the full range of appropriate years should be examined and reported” (italics original).

In our original analysis, we used self-esteem items from all waves and questionnaire forms from 1976 to 2006 and found that the correlation between self-esteem and year was −.019. We repeated the analysis restricting the focus to Form 5 data, which includes data from 1977 to 2006, and found the correlation was −.016. Thus, item order and context appear to exert little influence on the reported self-esteem trend. This similarity in effect sizes was not as pronounced, however, when looking at 2-year comparisons. That is, the $d$ for self-esteem comparing the mean from 1976 to the mean from 2006 using all forms was −.19, whereas the $d$ was −.02 when comparing 1977 to 2006 using only Form 5. We suspect that fluctuations in item ordering effects (when present) simply wash out over the longer span of time and increased sample size associated with using all available data. A similar difference is found in the 8th and 10th grade data reported by Twenge and Campbell (2010). They reported a $d$ of .12 for 8th graders and a $d$ of .06 for 10th graders and concluded that self-liking has increased from 1991 to 2007. Analyses using all waves of data showed that the correlation between year and self-liking for 8th and 10th graders was .03 and .01, respectively. These sorts of concerns are a major reason why we emphasized the full sample correlation coefficients in our original article, especially given that we found little evidence for curvilinear effects (as noted in the original).
Finally, Twenge and Campbell (2010) note that we only analyzed a subset of the items in the MTF project. We selected items that we thought were best matched to the Generation Me perspective that characterizes their description of today’s young people and we ultimately concluded that the evidence paints a far more subtle picture of generational differences.

Although we were limited by the number of constructs that could be examined given the space constraints of a single report, we had little explicit motivation to cherry pick results given that the data are available to academic researchers at many institutions through the ICPSR portal (www.icpsr.umich.edu). However, implicit biases are perhaps a fact of life, and we encourage other researchers to fully exploit the MTF and other existing datasets to evaluate strength of the evidence for generational changes for themselves.

Nonetheless, our interpretations of the existing data are also more or less consistent with interpretations offered by other researchers who have also looked at these issues. For instance, the editors of a recent special issue of the Journal of Managerial Psychology summarized the work with respect to the evidence and implications of generational differences at work: “[M]any of the empirical findings are less strong and consistent than popular sentiment suggests. Indeed, there may be more variation among members within a generation than there is between generations” (Macky, Gardner, & Forsyth, 2008, p. 860). Our conclusion is also consistent with an earlier report issued by the National Research Council (Sackett & Mavor, 2003). In this work, a panel of prominent scholars considered the implications of potential generational changes in youth attitudes and behaviors for military recruiting, including changes evident in the MTF project. The panel issued this summary:

“[C]ontrary to claims of large and dramatic differences among youth cohorts in different generations, high-quality longitudinal research documents a high degree of stability in
youth attitudes and values. Change is limited, and when it does occur, it occurs gradually. In addition, the popular literature is often based on selective, not systematic, data and analysis and on nonrepresentative samples. The committee does not believe that it is appropriate to give credence to popular portrayals of “generations” as a key explanatory concept for understanding youth attitudes and behaviors.” (pp. 305–306)

Acknowledgments—We thank Christopher Ferguson, Christopher Hopwood, Richard Lucas, and Frederick Oswald for valuable feedback on earlier versions and sections. We also thank Jerald Bachman and Patrick O’Malley. However, we take sole responsibility for the content and opinions expressed. Both authors contributed equally to this response.

References


Ferguson, C.J. (2008). The school shooting/violent video game link: Causal link or moral panic?


*Professional Psychology: Research and Practice.*


Roberts, B.W., Edmonds, G., & Grijalva, E. (2010). It is Developmental Me not Generation Me: Developmental changes are more important than generational changes in narcissism. *Perspectives on Psychological Science, 5*, xx–xx.


As we speculated in the original article, the maturity principle of personality development may itself play a role in the formation and persistence of negative stereotypes about young people.
When judging younger people, older adults may overlook the fact that their own personalities have matured as they assumed the “typical” tasks and responsibilities of adulthood, such as earning a living, participating in a committed romantic partnership, and raising children. Work by Richard Eibach and his colleagues (e.g., Eibach & Libby, 2008; Eibach, Libby, & Gilovich, 2003) provides an interesting explanation for the seemingly robust phenomenon in which older individuals criticize and pass moralistic judgments on the younger generation.

One last parenthetical comment on narcissism is worth noting. Twenge and Campbell (2010) refer to the seemingly high prevalence rates for Narcissistic Personality Disorder found in the National Epidemiologic Survey on Alcohol and Related Conditions. These findings are provocative, but the accuracy of prevalence estimates from that project has been questioned by some experts (e.g., Lenzenweger, 2008, p. 400; Lenzenweger, Lane, Loranger, & Kessler, 2007, p. 560). The original investigators did not use a validated Axis II measure of personality disorders and other studies using validated measures point to lower prevalence estimates for personality disorders (see Lenzenweger, 2008, p. 398).

It is also the case that rates of serious violent crimes by young people have similarly declined since the 1990s (see http://childstats.gov/americaschildren/beh5.asp).