Another Attempt to Move Beyond the Cross-Sectional U Shape of Happiness: A Reply

Nancy L. Galambos¹, Harvey J. Krahn², Matthew D. Johnson³, and Margie E. Lachman⁴

¹Department of Psychology, University of Alberta; ²Department of Sociology, University of Alberta; ³Department of Human Ecology, University of Alberta; and ⁴Department of Psychology, Brandeis University

Abstract

In this reply to Blanchflower and Graham (2021), we address key points in their response to our article, which challenged the robustness of the U shape in different contexts and aimed to move scholarship on life-course happiness beyond a focus on a single cross-sectional curve. We attend to apparent misconceptions in their portrayal of our arguments, identify points of agreement, and emphasize the value of exploring diversity rather than searching for homogeneity in patterns of change in well-being across life. Future longitudinal research to uncover the complex processes that characterize and underlie human happiness will lead to greater understanding that can benefit individuals and societies.

Keywords

happiness, life satisfaction, U shape, midlife crisis, life course

The main goal of our previous article published in Perspectives on Psychological Science (“The U Shape of Happiness Across the Life Course: Expanding the Discussion”; Galambos et al., 2020) was “to move beyond the disciplinary debate around whether mean levels of happiness are either higher or lower in midlife than in other age periods” (p. 899). We stated that “Instead, we should focus more on variability within and across people, highlighting the importance of questions about when and why some people are unhappy and what we can do about it” (p. 899). In that article we also made “several recommendations for steps that researchers can take to facilitate understanding across disciplines with an eye toward establishing fruitful multidisciplinary collaborations” (p. 908). In this reply to the Blanchflower and Graham (2020) response to our article (“The U Shape of Happiness: A Response”), we reiterate our position and emphasize our call for further inquiry.

Life-course (or life-span) research on adulthood is a vibrant field that addresses the complexities of intrapersonal development and individual and historical differences in patterns of change (e.g., Gerstorf et al., 2020; Inurna et al., 2020). There is a theory-rich literature in psychology on how well-being changes in adulthood (e.g., Mroczek & Spiro, 2005; Ryff, 1989) that presents a more nuanced view than is found in Blanchflower and Graham’s U shape approach. Guided by a life-course perspective, we suggested in our original article (Galambos et al., 2020) that researchers should explore variations in well-being in adulthood and their antecedents and consequences. Such an approach can enrich our understanding of why some adults can maintain or even increase their well-being throughout life, whereas others experience low points or ups and downs during adulthood. Moreover, this life-span developmental framework has important implications for designing interventions to prevent or treat declines in life satisfaction and optimize mental health throughout adulthood.

Addressing the Misconceptions

In their response to our article, Blanchflower and Graham state that the evidence they present “is overwhelming”...
(p. 1441) and “completely inconsistent with” our claims of diversity (p. 1441). They conclude that “There are U shapes in age” (p. 1441) in well-being. On that point we agree, and our original article cited some research in support of the U shape. Yet we hope to move beyond the findings from this one approach to deepen our understanding of well-being across the life span.

Blanchflower and Graham’s work on the U shape makes much of cross-sectional findings that, on average, those in their mid to late 40s or early 50s may show a low point in life satisfaction. Mean levels alone, however, do not provide important information about variability in happiness at any given age. Evidence suggests that only some people show a nadir in midlife. Bittmann (2021), for example, conducted person-centered analyses on cross-sectional data on life satisfaction across the adult life span in 81 countries and found three clusters of age-related patterns (downward, U shape, and downward with a stable period in late life), and the linear decline was evident in less developed countries.

Blanchflower and Graham say they are deeply concerned that we are “dismissive of phenomena that truly affect the lives of millions of people around the world” (p. 1435). We argue, however, that using such mean-level differences to portray midlife as a low point could be seen as irresponsible, promoting misrepresentations about a normative midlife crisis such as portrayed in the media (Lachman et al., 2015). Nevertheless, we agree with Blanchflower and Graham’s conclusion that findings from this area of inquiry merit “better understanding by both scholars and policymakers” (p. 1435). Indeed, this was precisely a key goal of our original article. As life-course scholars, we have devoted our careers to understanding developmental phenomena from adolescence through old age (e.g., Galambos et al., 2018; Infurna et al., 2020; Johnson et al., 2014; Krahn et al., 2018; Lachman, 2004; Lachman et al., 2015; Vargas et al. 2015), and we hope this exchange can help the field move beyond the cross-sectional U shape of happiness to explore processes of change in well-being.

In this reply to Blanchflower and Graham, we briefly review the six key points in our original article (Galambos et al., 2020), as it seems that they missed or ignored several of them, and we respond to some of their apparent misrepresentations and inaccuracies. We outlined the conceptual and methodological limitations of the argument that happiness (or well-being) consistently follows a U shape across the life span, with a low in midlife. We referenced previous studies on life satisfaction and happiness and reviewed a sample of 29 relevant empirical studies published in peer-reviewed journals between 2013 and 2019. It was not the goal of our article to do an exhaustive literature review or meta-analysis or to tally up the total number of studies that found or did not find the U shape. Instead, we wanted to show support for the view that not all researchers find the U shape, and when they do, they often also illustrate variability in age-related patterns of happiness. We concluded that the U shape is not as generalizable or robust as often portrayed and that to move the field forward, we need to recognize, study, and understand the great diversity and variation that exist in happiness trajectories (i.e., differential patterns of change) across the life course.

---

Cross-sectional studies are inadequate for drawing conclusions about within-person change in happiness across the life span

We cited conceptual and methodological support for our points that cross-sectional studies confound age with cohort differences and are not appropriate for drawing conclusions about within-person change because they assess only between-person (age) differences. Thus, because most studies supporting the U shape are cross-sectional, their contribution to understanding the development of well-being across the life course is limited.

Blanchflower and Graham are essentially silent on this fundamental issue, instead criticizing longitudinal studies because of attrition (p. 1436). Yet selective attrition in longitudinal research is something that can be identified, characterized, and addressed. On the other hand, cross-sectional methods also have design limitations. They not only confound age and cohort but also have selection issues. The pool of respondents at different ages for cross-sectional studies is inherently biased because, for example, those at older ages who have died are no longer available for inclusion in the study. We note that support for the U shape requires that both younger and older adults have higher levels of happiness than those in midlife. In their analysis of life satisfaction data using the Health and Retirement Study, Hudomiet et al. (2020) show that cross-sectional findings can bias the age profile upward because of differential mortality and nonresponse: “Individuals with higher life satisfaction and in better health tend to live longer and to remain in the survey, causing average values to increase” and they “conclude that the optimistic view about increasing life satisfaction at older ages based on cross-sectional data is not warranted” (p. 2).

Not only do Blanchflower and Graham offer no insight into why they think between-person comparisons are appropriate for drawing conclusions about within-person change in well-being across the life span, but
also they contend that their analyses with selected con-
trol variables enable them to discern “pure effects of aging” (p. 1434). The logic behind this statement is that if all potential confounds that could account for cohort differences are controlled, then the age differences in a cross-sectional analysis would reveal aging effects. Notwithstanding the remarkable assumptions underlying this argument (e.g., that all potential confounds have, in fact, been identified and controlled), life-span scholars would still disagree, arguing that, to examine aging, it is necessary to follow the same people over time (Baltes et al., 1977). There has been a long-standing discussion in the U-shape dialogue about what variables to control so that age differences can be more clearly identified (Morgan & O’Connor, 2020). Our key concern about cross-
sectional studies, however, is not what variables to control; it is whether cross-sectional studies, even those asking the same questions of different samples year after year (e.g., General Social Survey data), can provide insight into processes of aging.

By definition, aging is a complex set of interrelated processes involving biological, cognitive, psychological, and socioemotional changes taking place over time within—and not between—individuals (Baltes et al., 2006). Blanchflower and Graham’s focus on generating ever more cross-sectional studies does not bury the basic fact that aging is a within-person phenomenon best observed with repeated assessments of the same people using longitudinal designs. Interdisciplinary dia-
logues about the meaning and measurement of aging, and the research designs most able to capture processes of human development in motion, could move scholarship on happiness across the life course beyond a sin-
gular between-persons U-shape focus.

**Cross-sectional support for the U shape is mixed**

We cited previous literature questioning the universality of the U shape as promoted by Blanchflower and Graham and presented 20 cross-sectional studies in Table 1 (pp. 901–902 in Galambos et al.). To maintain comparability across studies, we intentionally focused our review on measures of subjective well-being most commonly used in the literature on age-related happiness (e.g., one-item life satisfaction), and we highlighted in Table 1 the measures used to draw our conclusions. We concluded that support for the U was not universal, and we explained our decisions.

In response, Blanchflower and Graham argued that we should not have concluded that “several articles did not find U shapes” (p. 1435) because quotes from authors of these articles show otherwise. We drew conclusions on the basis of a careful analysis of the results, and we stand by the validity of our coding of these studies as “yes” (support for the U), “no” (not support-
ive), or “mixed” (inconsistent support for the U across analyses). We also note that for every quotation that Blanchflower and Graham provide in support of the U for those studies we classified as mixed, there is another statement in the same article that recognizes diversity in the results. To give just one example, Laaksonen (2018) indeed stated that “we find some support for the U-shape curve over 30 countries” (p. 478), but they also concluded:

> the U-shape is not as simple as some research suggests, and thus the minimum happiness is not necessarily at approximately 40–50 years old. Minimum happiness can occur earlier or much later, depending on the model used, and the country concerned. The U-shape is clearly found in approximately one half of the 28 countries. A special feature is that the U-shape phenomenon holds better for males than for females. (p. 481)

Blanchflower and Graham also assert that they located many more papers purportedly supporting the U shape. As we already noted, our goal was not to conduct an exhaustive literature review. Instead our search for recent relevant articles was designed to inform our discussion of the issues, and we defined the parameters so that other scholars could replicate and expand it in scope. Without a narrative, system-
atic, or meta-analytic review published in a peer-
reviewed journal, however, it is not possible to judge their claim that all studies they located are supportive of their stance on the U shape. More importantly, we hold that what matters most for understanding happi-
ness across the life span is not the number of studies that support one position but the fact that there are a variety of outcomes and conclusions. It is also impor-
tant to understand what the cross-sectional findings mean.

Since publishing our article, we have become aware of additional relevant studies. Morgan and O’Connor (2017), for example, reported an M shape in Euroba-
rometer life-satisfaction data. More recently, they argued that “ . . . the U-shaped relation is, in fact, not every-
where” (Morgan & O’Connor, 2020, p. 201). Likewise, using Gallup data from 166 countries, Jebb et al. (2020) documented age differences so small that they con-
cluded “much about the U shape has been overblown” (p. 302). Blanchflower and Graham dismiss Jebb et al.’s conclusion as being “without foundation” (p. 1440), and present a graph on age differences in life satisfaction in recent data from the United Kingdom (Figure 1 in Blanchflower and Graham). They urge the reader to
Thus, we examined that graph (reproduced here as Fig. 1) and noticed that Blanchflower and Graham severely truncated the y-axis with a range from 7.0 to 8.5 on an 11-point scale, despite our earlier recommendation that full response scales should be shown to visually portray the magnitude of any age differences. On that truncated scale, the pattern looks like a U. When replotted to represent a larger portion of the distribution that includes the range of likely responses on the scale (5–10; Fig. 2) or the full scale (0–10; Fig. 3), the data look more like a straight line. Such truncated depictions are not uncommon in presentations of data in support of the U shape (see e.g., Figures 2–4 in Blanchflower and Graham), but if most of a scale has to be trimmed to illustrate a U, then Jebb et al. (2020) are correct in their assertion that the age trends are small. How consequential is a difference between 7.9 and 7.2 on an 11-point scale, particularly when this range illustrates quite high life satisfaction above the midpoint? Equally critical, such graphs should include error bars to show the likely overlap in life satisfaction across ages and which age differences are significant (see e.g., Helliwell et al., 2019).

**Longitudinal support for the U shape is mixed**

We examined the sparse literature on longitudinal and accelerated longitudinal studies of the age-happiness association and identified 13 relevant studies that again showed a lack of universal support consistent with a U shape. Blanchflower and Graham make a puzzling series of points to contest our assessment of studies finding no or mixed support.

First, Blanchflower and Graham claim that several longitudinal studies should have been eliminated from our analysis because they did not meet our criteria for inclusion in our review. Our explicitly stated inclusion criterion for longitudinal studies was that they “had at least two times of measurement and spanned the periods of young adulthood to midlife, midlife to late life, or young adulthood to midlife to late life” (Galambos et al., 2020, p. 900). Blanchflower and Graham inappropriately applied our inclusion criterion for cross-sectional studies (“spanned the teens or 20s into the 60s”) to longitudinal studies that challenged the U shape.

Second, Blanchflower and Graham claim our summary of the Cheng et al. (2017) article “misrepresents what they [Cheng et al.] said” (Blanchflower & Graham, 2021, p. 1437), and they enlist the second author on that study (N. Powdthavee) to repeat their allegation. We invite readers to check our statements (p. 904 in Galambos et al.) against the original text of the Cheng et al. article (p. 127). The opening of the Cheng et al. article states: “There is a large amount of cross-sectional evidence for a midlife low in the life cycle of human happiness and well-being (a ‘U shape’). Yet no genuinely longitudinal inquiry has uncovered evidence for a U-shaped pattern” (p. 126). We know Cheng et al. used this as a rationale for their study and presented their
longitudinal findings as an exception. We specifically acknowledged their research as supportive of the U curve: twice in Table 1 (pp. 901–902) and twice in the text (pp. 904–905). We did not take their “findings entirely out of context” (Blanchflower and Graham, 2021, p. 1437), and we stand by the accuracy of our representation of how Cheng et al. characterized the state of previous longitudinal research.

Third, Blanchflower and Graham supply selective quotations to reject our conclusions for longitudinal...
studies with mixed or no support for the U shape. For example, they quote Bauer et al. (2017) with respect to analyses supporting the U, but neglect to mention Bauer et al.’s telling conclusion: “While the U-shape remains stable in Western Europe, well-being seems to mostly decline over the lifespan in the ECA countries” (p. 146). It is noteworthy that scholars like Bauer et al. look for and find heterogeneity in the age-happiness connection. Indeed, using cross-sectional Gallup Poll data, Helliwell et al. (2019) recently found diversity in patterns of age differences in life satisfaction across nine regions of the world. Although life satisfaction was lower in midlife than in young adulthood in most global regions, this trend was flatter for some areas than others, and only two regions showed a clear upward trend after midlife (North America/Australia/New Zealand and East Asia). We applaud such efforts as they advance research by delving deeper into the mysteries of the complex connections between human aging and well-being. Similar explorations of heterogeneity in longitudinal studies would advance the goal of learning more about variation in the developmental course of well-being.

Fourth, Blanchflower and Graham say “Sample sizes below 2,000 and even more so those below 1,000 just do not have enough explanatory power to generate statistically significant econometric results” (p. 1435). We must ask: If an effect size is so small as to require samples in the many thousands to attain statistical significance, how meaningful are the age differences? Many highly influential studies in psychology and sociology have presented statistically significant findings with samples containing fewer than 1,000 participants. Furthermore, we point out that longitudinal studies such as ours (Galambos et al., 2015) have multiple observations per participant. We analyzed longitudinal data for two cohorts, one surveyed six times over 14 years (about 2,500 person-year data points) and the other surveyed seven times over 25 years (about 4,000 person-year data points), and we reported statistically significant results. In fact, Rast and Hofer (2014) documented that most developmental longitudinal studies have ample power to detect change over time, especially when there are many waves over a long duration (e.g., decades).

**Longitudinal research on other well-being indicators challenges the U shape**

It is important to recognize that midlife is not the only age period in which there can be distress. Understanding the state of well-being among young people is critical, given that two thirds of Canadians under age 25 recently reported symptoms consistent with generalized anxiety disorder and/or major depressive disorder (Nwachukwu et al., 2020), and the prevalence of mood disorders and suicidal outcomes among 18- to 25-year-olds in the United States was substantially higher in 2015 through 2017 than for that age group in the period from 2008 through 2014 (Twenge et al., 2019). These alarming trends can be missed with a focus on the U shape. We provided evidence from longitudinal studies on various mental-health indicators (e.g., depression, anger), showing that mental health actually improves rather than falls from adolescence into midlife. Blanchflower and Graham are silent on the meaning of the longitudinal evidence we presented and do not comment on the implications of studies demonstrating significant heterogeneity in trajectories of change in different indicators of well-being across life. Their response is to present cross-sectional results on mental-health indicators without providing a compelling rationale for why their studies of age differences should be interpreted as evidence of age-related change.

Moreover, they cite the cross-sectional National Comorbidity Survey Replication (Kessler et al., 2010) as showing midlife peaks in a variety of mental-health problems, including major depressive episodes, yet a close look at the results reveals that for many dimensions, the oldest age group (≥ 65 years old) showed considerably lower prevalence of problems than all the younger age groups (18–34, 35–49, 50–64) who were similar to each other. Because the U-curve depends on midlifers being less happy than young adults, we disagree that these data constitute support for a low point in well-being that is unique to midlife. On this point we agree: Employing multiple indicators of well-being is useful for gaining a more complete understanding of how people might change with age.

**Reflections on the past highlight midlife as happy**

We presented data from multiple studies showing that many people, when looking back to midlife, recall it as happy or one of the happier periods of life, especially compared with adolescence. This is important because U-curve proponents view adolescence and young adulthood as happier periods than midlife. Blanchflower and Graham dismiss this line of research because of the limitations of retrospective recall and small sample sizes, referring to the evidence we provided as anecdotes from “elderly” respondents rather than accurately characterizing the wide age ranges of respondents in several studies we cited, including our own, in which 50-year-olds were asked to reflect on their best decade of life.
Furthermore, Blanchflower and Graham argued that factors such as illness and poverty would differentially shape the recall of individuals in such studies. Ironically, this very argument is key to our focus on diversity in paths of happiness through life. Many factors (e.g., illness, poverty, immigration, peacetime vs. wartime) are sources of diversity in happiness trajectories. They are not to be explained away or controlled but to be considered as potential predictors so that we can better understand interindividual differences in intraindividual change in happiness.

Blanchflower and Graham say the results of recall studies are likely the consequence of present bias. This is indeed a limitation of retrospective studies, but one must acknowledge that all self-report measures raise similar concerns. Just as subsequent experiences can influence one's recall of earlier times, the kind of day or week one is having can affect concurrent responses on survey questions about life satisfaction or happiness. Although how one reconstructs the past is relevant for mental health in and of itself, prospective studies can help to examine the accuracy of retrospective accounts. One study, for example, found that retrospective accounts of life satisfaction 10 years earlier were more accurate for older than for younger adults (Lachman et al., 2008). And we showed (Galambos et al., 2020, Figure 1 and p. 906) that retrospective reports on the best decade of life mapped quite well onto reports of happiness that were assessed in adolescence, young adulthood, and midlife. A triangulation of methods is the best antidote to the biases that can result from any one approach. Retrospective accounts can be used to elaborate on experiences across the life span rather than relying on a single curve to characterize human development.

The most important question concerns diversity in happiness and its sources

We argued for renewed emphasis on diversity, stating that “a focus on a single trajectory of well-being is of limited scientific and applied value because it obscures the diversity in pathways throughout life as well as its sources” (p. 898). We used the term “trajectory” as it is typically used, to refer to a path, curve, or progression of development, and our argument was directed toward the limitations of ignoring interindividual differences (diversity) in intraindividual changes in happiness. Blanchflower and Graham appear to have misunderstood, because they responded by insisting the use of one-item measures of well-being is acceptable. This is not a point of contention; although we recommend the use of multiple and multi-item measures when possible, we also support the use of one-item measures.

But Blanchflower and Graham provide no response as to the importance of considering diversity in patterns. They only repeat their claim that “the evidence for a U shape in well-being is overwhelming” (p. 1441). If one looks at longitudinal changes, however, one can easily see that some people are stable, some increase, and some show a decrease in midlife (Röcke & Lachman, 2008). Our central point is that one curve of means does not represent the range of possible patterns of change and masks individual differences. A more person-centered approach is needed to characterize variations in patterns of change. Indeed, longitudinal work suggests that most people are relatively stable in life satisfaction across adulthood (Lachman et al., 2008). Attending to whom is unhappy when, where, and why, is more likely to lead to targeted prevention and intervention efforts that will help those most in need rather than focusing on one age group purported to be the unhappiest.

Conclusion

We conclude our reply to Blanchflower and Graham by quoting Kiri et al. (2018):

not only are scientific findings never complete or definitive and are always prone to improvement; but, also, that observing only apparently definitive or undisputed findings may be a sign of weakness of a scientific field rather than a proof of its solidity. (p. 835)

It is certainly the case that reproducibility in science is both desirable and important. Blanchflower and Graham and their colleagues have replicated their analyses and findings on the U shape over and over again using the same methods. But they discount the ample evidence of deviations from the U shape seen in data across the world and published by other authors in multiple disciplines. To move science forward and advance our understanding of a phenomenon, we believe a triangulation of methods is called for, to ask different questions, using a variety of measures, designs, samples, and analytic techniques. In the end, we stand by our original conclusion: “Ultimately, it would be ideal if the results of research on happiness and well-being over the life course could be applied to understanding the variations within age groups with an eye toward addressing the sources and consequences of unhappiness” (p. 908). Such understanding could inform personalized treatments and interventions to address issues of mental health in adulthood. It is time to move beyond a preoccupation with the shape of cross-sectional data on life satisfaction and happiness.
Transparency

Declaration of Conflicting Interests

Funding

This work was supported by the Social Sciences and Humanities Research Council of Canada (Grant 435-2014-0076) and the National Institute on Aging (Grants P01-AG20166 and U19-AG051426).

References


