HHS Public Access

Author manuscript

Psychol Aging. Author manuscript; available in PMC 2016 March 01.

Published in final edited form as:

Psychol Aging. 2015 March ; 30(1): 106–119. doi:10.1037/a0038502.

Age Differences and Longitudinal Change in the Effects of Data Collection Mode on Self-Reports of Psychosocial Functioning

Gloria Luong,
Max Planck Institute for Human Development

Susan T. Charles,
University of California, Irvine

Karen S. Rook,
University of California, Irvine

Chandra A. Reynolds, and
University of California, Riverside

Margaret Gatz
University of Southern California and the Karolinska Institutet

Abstract

The current study investigated how participant aging may influence *mode effects*, wherein individuals report less negative and more positive psychosocial functioning with data collection modes that have greater (vs. less) direct contact with interviewers (e.g., in-person interviews vs. telephone interviews). Using two longitudinal datasets, the Later Life Study of Social Exchanges (LLSSE) and Swedish Adoption/Twin Study of Aging (SATSA), we tested how mode effects may vary with cohort (baseline age differences) and maturational development (longitudinal change). In Study 1, LLSSE participants (65–90 years old) completed in-person and telephone interviews assessing negative and positive aspects of psychosocial functioning across two years. The data collection mode with greater direct contact with interviewers (in-person interviews) was associated with reporting less negative and more positive psychosocial functioning compared to the mode with less direct contact (telephone interviews). These mode effects were more pronounced with older baseline age, but only for the negative psychosocial measures. Mode effects also became stronger over time for reports of negative affect. In Study 2, SATSA participants (38–86 years old) completed mailed questionnaires and questionnaires collected in-person that assessed depressive symptoms and positive affect across 18 years. Consistent with Study 1, participants reported fewer depressive symptoms and more positive affect with greater (vs. less) direct contact with interviewers (questionnaires collected in-person vs. mailed questionnaires). For reports of depressive symptoms, but not positive affect, mode effects were

Correspondence should be addressed to Gloria Luong, Max Planck Research Group “Affect Across the Lifespan,” Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany. Phone: +49 30 82406 342. Fax: +49 30 82499 39. Luong@mpib-berlin.mpg.de.
Gloria Luong, Max Planck Research Group “Affect Across the Lifespan,” Max Planck Institute for Human Development; Susan T. Charles and Karen S. Rook, Department of Psychology and Social Behavior, University of California, Irvine; Chandra A. Reynolds, Department of Psychology, University of California, Riverside; Margaret Gatz, Department of Psychology, University of Southern California and Karolinska Institutet, Stockholm, Sweden.

more pronounced with age and time. Together, the results underscore how mode effects may contribute to inconsistent findings in the socioemotional aging literature.

Keywords

Data collection method effects; social desirability; affective well-being; socioemotional development; depression

Rapid population aging projections in developed nations (see review by Christensen, Doblhammer, Rau, & Vaupel, 2009) have motivated scientists to better understand age-related changes in psychosocial functioning (e.g., depression, affective well-being, social interactions). Accordingly, predominant research approaches rely on participant self-reports. Although researchers have long recognized *mode effects*, wherein data collection modes with greater (vs. less) direct contact with interviewers elicit less negative and/or more positive (desirable) self-reports from participants, these effects have often been overlooked in the socioemotional aging literature (see review by Richman, Kiesler, Weisband, & Drasgow, 1999). The mere existence of age-related changes in mode effects, however, would mean that studies employing different data collection modes can obtain different trajectories of socioemotional aging (e.g., blunted decline or more favorable functioning with age using modes that involve greater direct contact with interviewers). Age differences in mode effects may therefore be contributing to inconsistent results across studies in the socioemotional aging literature. Thus, the main purpose of the current study was to investigate a methodological question regarding whether age differences and longitudinal age changes exist in mode effects on self-reports of psychosocial functioning, and if so, to consider its implications when interpreting the socioemotional development literature.

Effects of Data Collection Mode on Participant Self-Reports of Psychosocial Functioning

Much of the literature on psychosocial functioning is based on self-reported data. Although self-reports can provide insights into participants' experiences, they may also be susceptible to response biases. One type of response bias, referred to as mode effects, arises when participants report more favorable psychosocial functioning with data collection methods that involve greater (vs. less) direct interviewer contact, such as when comparing in-person interviews to telephone interviews. For example, when in greater direct contact with interviewers, participants report less negative outcomes, such as less severe injuries (Hoher, Bach, Munster, Bouillon, & Tiling, 1997), less hazardous drinking behaviors (Gmel, 2000), and are less likely to report an abortion and illicit drug use (see review by Tourangeau & Smith, 1996), compared to other individuals who answer the same questions with less direct interviewer contact. These mode effects also apply to the reporting of more positive psychosocial functioning, such as church attendance (Presser & Stinson, 1998) and voting behaviors (Voogt & Saris, 2005). General response biases are attributed to participants': (a) *abilities* (e.g., level of introspection, cognitive capacities) and/or (b) *willingness* to report on psychosocial experiences (e.g., frankness, motivation to disclose personal information; cf. Nunnally, 1978). Although both of these factors may vary across modes and contribute to

different reports across modes, researchers have argued that mode effects are primarily due to participants' reduced *willingness* to report poor functioning in modes with greater direct interviewer contact, given that these modes entail a stronger risk of (negative) personal evaluation and higher demand characteristics (e.g., Bowling, 2005; Kreuter, Presser, & Tourangeau, 2008).

The Role of Participant Aging on Mode Effects

The available evidence suggests that participant willingness to report poor psychosocial functioning may decrease with age, and this effect may be exacerbated with data collection modes involving greater direct interviewer contact. Three inter-related factors may contribute to age differences in mode effects: (a) cohort differences in willingness to disclose poor psychosocial functioning, (b) age-related increases in the prioritization of socioemotional goals, and (c) age-related increases in socially desirable responding.

First, there may be age differences in mode effects given that individuals born in Westernized countries in the early twentieth century were socialized to "keep a stiff upper lip" and refrain from sharing social and emotional problems (Bennett, 2007). These cohort effects would therefore predict that older individuals are generally less willing to report poor psychosocial functioning. Additionally, age differences in mode effects may be partly influenced by developmental changes wherein socioemotional concerns, such as maintaining high levels of emotional well-being and social harmony, take on greater importance across adulthood and become central to older adults' motivational strivings (Carstensen, Isaacowitz, & Charles, 1999; Riediger, Schmiedek, Wagner, & Lindenberger, 2009). Although it may be unpleasant for participants to report poor psychosocial functioning, older individuals who highly value socioemotional well-being may find it particularly stigmatizing to admit they are failing to attain these important personal goals, compared to younger individuals, for whom such goals may have lower priority. As discussed previously, greater direct interviewer contact entails stronger potential for participants to be negatively evaluated. Thus, reporting poor psychosocial functioning to an interviewer may further elicit low self-esteem and personal self-regard, as well as greater negative affect (e.g., embarrassment), thereby undermining the very socioemotional goals that take precedence in later life. Older individuals may therefore be highly motivated to report more favorable psychosocial functioning, given that such reports may help them save face and dampen negative emotional responses to potentially poor evaluations by the interviewer (e.g., Dijkstra, Smit, & Comijs, 2001; Ray, 1988; Unützer, 2002).

Indeed, when older adults perceive that they may be judged negatively on core aspects of the self and important personal goals (socioemotional well-being), they are more likely to provide positively-biased reports of their psychosocial functioning compared to younger individuals. For example, when older adults are asked a *personally evaluative* question about their abilities to handle their current circumstances (a question which involves potential for participants to be negatively judged by the researchers on their coping potential), they report more positively-biased responses than when they are simply asked to describe their circumstances (a question which involves less potential for negative personal evaluation; Carp & Carp, 1981). Finally, many studies have shown that older age is

associated with more socially desirable responding (i.e., less negative and/or more positive outcomes) on measures of depression, mood, and personality (Carstensen & Cone, 1983; Lewinsohn, Seeley, Roberts, & Allen, 1997; Thomsen, Mehlsen, Viidik, Sommerlund, & Zachariae, 2005). Together, the literature points to a general tendency by older adults to report more favorable psychosocial functioning compared to their younger counterparts, which may be enhanced with greater direct interviewer contact. Thus, mode effects may be more pronounced with age.

Implications of Mode Effects for Interpreting Aging Effects in Psychosocial Functioning

The existence of age differences in mode effects would have important implications not only for the varying pattern of results that researchers may obtain by using different data collection modes across studies, but also for researchers' interpretation of these results. For example, accumulating research suggests psychosocial adjustment improves across adulthood, with age-related decreases in depressive symptoms (e.g., Jorm, 2000; Kessler et al., 2010), negative affect (e.g., Carstensen et al., 2011; Charles, Reynolds, & Gatz, 2001; Riediger et al., 2009), and negative social exchanges (Birditt, Fingerman, & Almeida, 2005), as well as stability or increases in positive affect (Charles et al., 2001; Mrozcek & Kolarz, 1998) and satisfaction with one's social network (Lansford, Sherman, & Antonucci, 1998; see review by Luong, Fingerman, & Charles, 2011). These patterns are not ubiquitous, however, and some studies find contrasting results, such as upturns in depressive symptoms (Gatz & Hurwicz, 1990; Pinquart, 2001; Snowdon, 2001; Stordal et al., 2001) and sadness (Kunzmann, Richter, & Schmukle, 2013) in later life. Other studies have shown that age is not directly related to satisfaction with one's social life (von Hippel, Henry, & Matovic, 2008). Although mode effects may be contributing to some of the discrepancies between studies, they have often been ignored in the socioemotional aging literature. As discussed, however, there may indeed be age differences in mode effects. In fact, studies have shown that when adjusting for older adults' positively-biased response tendencies, age differences in psychosocial functioning are attenuated (e.g., Soubelet & Salthouse, 2011). Thus, older adults' greater tendency to provide more favorable responses in modes with greater direct contact with interviewers may be contributing to inconsistent patterns of results across studies.

To illustrate this point, in the depression literature, estimates of the prevalence of late-life depression vary widely across studies, ranging from as little as 3% to as much as over 38% (Murrell, Himmelfarb, & Wright, 1983). In some meta-analyses, literature reviews, and empirical studies that aggregate results across data collection methods, there appear to be no clear age-related patterns in depressive symptoms (e.g., Jorm, 2000; Spinhoven et al., 1997). Within the last decade or so, researchers have suggested that inconsistencies in patterns of age-related differences in psychosocial adjustment across studies may be partly accounted for by mode effects (e.g., Beekman, Copeland, & Prince, 1999; Jorm, 2000; Stordal et al., 2001). Specifically, studies using data collection modes that involve greater direct interviewer contact (e.g., in-person interviews) tend to find age-related decreases in reports of depressive symptoms (e.g., Eaton & Kessler, 1981; Henderson et al., 1998), whereas

studies that involve less direct contact (e.g., questionnaires) may be more likely to show age-related increases (e.g., Gatz & Hurwicz, 1990; Stordal et al., 2001). Thus, if certain modes are more likely to accentuate (or attenuate) age differences in reports of psychosocial functioning, it is crucial to understand how mode effects may generate discrepancies in the literature.

The Current Study

Despite the important implications of mode effects, few studies have directly and systematically investigated how these effects are moderated by participant aging. The few studies that have examined age differences in mode effects have suggested that, under certain conditions, older individuals may be more strongly influenced by mode effects compared to younger individuals (e.g., Evans, Kessler, Lewis, Peters, & Sharp, 2004; Herman, 1977; Herzog & Rodgers, 1988; Lang, John, Lüdtke, Schupp, & Wagner, 2011). Many of these studies, however, utilized a between-subjects design, whereby responses from participants who answer questions using one type of mode (e.g., in-person interviews) are contrasted with those of different participants answering the same questions using another mode (e.g., telephone interviews). Between-subjects designs make it nearly impossible to discern whether differential reports of socioemotional functioning are due to mode effects or selection effects, whereby certain types of individuals may be more likely to self-select into particular data collection modes (Vannieuwenhuyze, Loosveldt, & Molenberghs, 2010). Even when participants are randomly assigned to different modes, pre-existing group differences that may not be evident at baseline may emerge over time (e.g., Presser & Stinson, 1998). Longitudinal studies that administer multiple data collection modes to the same participants (i.e., within-subject designs) are ideal for detecting mode effects and changes in such effects, given that individuals serve as their own controls. Although some previous studies have used a longitudinal within-subjects approach, they have focused on other methodological issues, such as participant nonresponse (Schräpler et al., 2010) or measurement invariance (e.g., Lang et al., 2011), instead of variations in reports of psychosocial functioning by mode and age-related changes in such mode effects.

The current study builds on the literature by examining whether: (a) mode effects exist for widely used measures of socioemotional functioning (i.e., depressive symptoms, positive and negative affect, positive and negative social exchanges), and (b) these mode effects are moderated by participant age (cross-sectional baseline age) and aging (longitudinal change). We also explored whether these effects persist for measures of both negative and positive psychosocial adjustment. We predicted mode effects in both positive and negative psychosocial functioning, such that individuals would report less negative and more positive psychosocial adjustment with modes involving greater (vs. less) direct contact with interviewers. Moreover, we predicted that mode effects would be more pronounced with older age (at baseline and longitudinally over time). To test these hypotheses, we analyzed data from two longitudinal mixed-method studies (the Later Life Study of Social Exchanges (LLSSE) and Swedish Adoption/Twin Study of Aging (SATSA)) in which participants reported on a variety of psychosocial functioning measures via different data collection methods that varied in level of direct contact with interviewers over multiple occasions.

Study 1

Study 1 examines mode effects such that participants are predicted to report less negative (i.e., depressive symptoms, negative affect, negative social exchanges) and more positive (i.e., positive affect, positive social exchanges) psychosocial functioning with the data collection mode with greater direct contact with interviewers (in-person interviews (IPI)), compared to the mode with less direct contact (telephone interviews (TI)).

Method

Participants

Participants in the Later Life Study of Social Exchanges (LLSSE) were selected randomly from the Medicare Beneficiary Eligibility List of the Centers for Medicare and Medicaid Services (Sorkin & Rook, 2004). The sample was representative of the older adult population in the United States at the start of the study (87.1% White, 10.7% Black or African American, and 2.2% other (e.g., Asian American)). Of the 916 original participants, 788 were included in the present study because they completed at least one telephone interview in addition to the baseline in-person interview[1]. The sample averaged 73.89 years old (SD = 6.60, range 65–90 years) at baseline and included more women than men (62.6% vs. 37.4%). In this sample, 25.4% of the participants had attained less than a high school education, 35.8% had a high school degree, 19.8% had some college education, and 18.9% had a college degree or higher. Median pretax income at the baseline assessment was $14,000-$24,999.

Procedure

The current study used data from the five waves of the LLSSE. Data collection occurred every six months for two years and assessed a broad array of psychosocial indices. The data collection mode alternated between waves such that participants completed in-person interviews (IPI) for the baseline, third, and fifth waves of the study, and abbreviated telephone interviews (TI) for the second and fourth waves of the study. The IPI were approximately 70 minutes long and TI were approximately 15–30 minutes long.

Measures

Demographic characteristics—Demographic characteristics assessed included participants' age (birthdate), gender (effect coded: −1 = *male*, 1 = *female*), educational attainment, and marital status.

[1]Participants who completed at least one telephone interview in the LLSSE (and thus included in the current analysis) were significantly younger ($M(SD)_{\text{current sample}}$ = 73.89(6.60) years old vs. $M(SD)_{\text{excluded participants}}$ = 75.83(6.62) years old; $t(914) = -3.09$, $p = .002$), better educated ($M(SD)_{\text{current sample}}$ = 4.65(2.01) vs. $M(SD)_{\text{excluded participants}}$ = 3.97(1.74); $t(182.82) = 4.01$, $p < .001$, unequal variances), and reported fewer health conditions ($M(SD)_{\text{current sample}}$ = 2.15(1.48) vs. $M(SD)_{\text{excluded participants}}$ = 2.50(1.89); $t(153.53) = -2.02$, $p = .046$, unequal variances) and depressive symptoms ($M(SD)_{\text{current sample}}$ = 4.87(4.70) vs. $M(SD)_{\text{excluded participants}}$ = 6.55(6.32); $t(146.66) = -2.84$, $p = .005$, unequal variances) at the baseline assessment compared to participants who did not complete any telephone interviews. Participants in the current sample, however, did not differ from the excluded participants on their baseline reports of negative affect, $t(152.57) = -1.61$, $p = .11$, unequal variances, or negative social exchanges, $t(870) = 1.17$, $p = .24$.

Health conditions—Total number of health conditions was derived by summing across participants' reports of having been diagnosed with any of 13 health conditions, such as diabetes, asthma, and hypertension at the baseline assessment.

Negative Psychosocial Functioning

Depressive symptoms—Participants rated how often each of 11 different statements was true during the past week on a four-point scale from 0 (*none of the time*) to 3 (*most of the time*) as a measure of depressive symptoms. The items were incorporated from two previously validated short-form versions of the Center for Epidemiological Studies Depression Scale (CES-D; Santor & Coyne, 1997; Turvey, Wallace, & Herzog, 1999). We excluded 2 items related to positive affect such that scores were based on a total of 9 depressive symptoms items. Scores at each wave were based on the sum of the items.

Negative affect—Negative affect was assessed with five items (worried, frustrated, angry, unhappy, blue) on a 5-point scale from 0 (*very slightly or not at all*) to 4 (*very much*) during the last month (based on a validated measure developed by Diener & Emmons, 1984, Studies 3–5).

Negative social exchanges—Participants rated how often they had experienced negative social exchanges across 4 domains (unsympathetic/insensitive behaviors, unwanted advice/intrusion, failures to help/provide needs, and rejection/neglect) within the last month. Each domain was assessed by 3 items using a 5-point scale (0 = *never* to 4 = *very often*). The 12 items across the 4 domains were averaged together for a measure of frequency of negative social exchanges (see Newsom, Rook, Nishishiba, Sorkin, & Mahan, 2005 for more information).

Positive Psychosocial Functioning

Positive affect—Positive affect was assessed with five items (happy, satisfied, enjoy yourself, pleased, and joyful) on a 5-point scale from 0 (*very slightly or not at all*) to 4 (*very much*) during the last month (Diener & Emmons, 1984).

Positive social exchanges—Similar to the negative social exchanges measure, participants rated how often they had experienced positive social exchanges across 4 domains (emotional support, informational support, instrumental support, and companionship) within the past month, with each domain assessed with 3 items on a 5-point scale (0 = *never* to 4 = *very often*). The 12 items were averaged together.

Results

Descriptive Statistics and Preliminary Results

The means and internal consistencies for the three negative (i.e., depressive symptoms, negative affect, negative social exchanges) and two positive psychosocial outcomes (i.e., positive affect, positive social exchanges) at each wave are listed in Table 1a and Table 1b, respectively. Cronbach's alphas show high and stable internal consistencies for each psychosocial measure across waves and data collection modes. We next examined

associations between demographic characteristics (gender, age and education), health conditions, and the variables of interest in the study at baseline. Depressive symptoms were related to female gender ($r = .10$, $p < .01$), lower educational attainment ($r = -0.11$, $p < .01$), and more health conditions ($r = .25$, $p < .001$). Negative affect was linked to female gender ($r = .12$, $p < .01$) and more health conditions ($r = .18$, $p < .001$). Negative social exchanges were associated with younger age ($r = -0.17$, $p < .001$) and more health conditions ($r = .12$, $p < .01$. Because gender, education level, and baseline health conditions were related to the dependent variables, they were included as covariates in all subsequent analyses.

Overview of Data Analysis

To test for mode effects, we used multi-level modeling in SAS PROC MIXED v. 9.2 to account for the nested data. We used an unstructured covariance matrix with full maximum likelihood estimation methods. A two-level model was specified, whereby intra-individual changes in psychosocial reports across modes could be examined at Level 1 and inter-individual (between-subjects) effects testing baseline age differences at Level 2. Time and mode were specified as random slopes in the following models.

Mode Effects in Self-Reported Negative and Positive Psychosocial Functioning

First, we hypothesized that participants would report less negative psychosocial functioning (i.e., depressive symptoms, negative affect, and negative social exchanges) in the in-person interviews (IPI) compared to the telephone interviews (TI). We examined mode effects (0 = TI, 1 = IPI) at Level 1 for each of the psychosocial outcomes. In separate models, we tested for main effects of mode adjusting for baseline age, time (centered on the first wave), and the aforementioned covariates. For all three negative psychosocial outcomes, participants reported significantly lower scores when interviewed in-person compared to over the telephone (depressive symptoms: $B = -1.84$, $SE = .13$, $p < .001$, accounting for 23% of the explainable within-person variance in depressive symptoms; negative affect: $B = -0.15$, $SE = .02$, $p < .001$, accounting for 13% of the explainable within-person variance in negative affect; negative social exchanges: $B = -0.38$, $SE = .02$, $p < .001$, accounting for 37% of the explainable within-person variance in negative social exchanges[2]).

Similarly, we expected mode effects for positive psychosocial functioning, such that participants would report more positive affect and frequent positive social exchanges on the IPI relative to the TI. Consistent with our hypotheses, participants reported significantly greater positive affect, $B = 0.06$, $SE = .02$, $p < .001$ (accounting for 5% of the explainable within-person variance in positive affect), and more positive social exchanges, $B = .06$, $SE = .02$, $p < .01$ (accounting for 16% of the explainable within-person variance in positive social exchanges), on the IPI compared to the TI[2].

[2]Proportion of explainable variance was calculated as recommended by Singer (1998). The amount of explainable variance was based on a model with all covariates and baseline age (data collection mode was excluded). In the next model, mode was included. The proportion of explainable variance accounted for by mode was calculated as the difference in the residual variance between the two models (i.e., model without and with mode) divided by the total explainable variance.

Tests of Age-Related Moderation of Mode Effects in Self-Reported Negative and Positive Psychosocial Functioning

Negative psychosocial functioning—Next, we tested whether participant aging moderates mode effects in reports of negative psychosocial adjustment. We conceptualized participant aging by distinguishing between (a) cross-sectional baseline age differences in mode effects (suggestive of cohort influences) and (b) longitudinal changes in mode effects (suggestive of maturational development). We predicted that mode effects are greater with older baseline age and become more pronounced over time. We tested cross-product interactions between grand-mean centered baseline age and mode, as well as with time and mode in a full model that included the lower-ordered terms and covariates in the model. The Age × Mode interactions significantly predicted all three negative psychosocial measures (depressive symptoms, negative affect, and negative social exchanges) in separate models (see Table 2). Consistent with our predictions, the interactions revealed that even in this late life sample, increasingly older baseline age was related to stronger mode effects (i.e., larger differences in reports of depressive symptoms (Figure 1A), negative affect (Figure 1B), and negative social exchanges (Figure 1C) between modes). Additionally, the Time × Mode interaction was significant, but only for the negative affect model (see Table 2). In line with our predictions, however, the interaction showed that mode effects in reports of negative affect became stronger over time (Figure 1D).

Follow-up simple slope analyses were conducted to assist with interpretations of the Age × Mode interactions. Overall, these results indicated that the mode with greater direct interviewer contact (IPI) produced more favorable age-related estimates (i.e., no age-related increases in depressive symptoms, $t = 0.36$, $p = .72$; age-related decreases in both negative affect, $t = -2.40$, $p = .016$, and negative social exchanges, $t = -5.60$, $p < .001$) compared to the mode with less direct interviewer contact (TI; i.e., age-related increases in depressive symptoms, $t = 2.92$, $p = .004$; no age-related decreases in negative affect, $t = -0.55$, $p = .59$, or negative social exchanges, $t = -0.97$, $p = .33$).

Positive psychosocial functioning

We used the same approach to test whether participant aging moderates mode effects in reports of positive psychosocial functioning (i.e., positive affect, positive social exchanges). The results, shown in Table 2, revealed that unlike the models for negative psychosocial functioning, none of the interactions reached statistical significance.

Discussion

In the LLSSE, we found evidence for a main effect of mode such that participants reported more desirable (i.e., less negative, more positive) psychosocial functioning in the mode with greater direct interviewer contact (i.e., in-person interviews (IPI)) compared to the mode with less direct contact (i.e., telephone interviews (TI)). For the negative psychosocial measures (i.e., depressive symptoms, negative affect, negative social exchanges), the mode effects were stronger with older baseline age, even within this elderly sample. These interactions revealed that the IPI yielded more favorable age-related estimates of negative psychosocial functioning (e.g., no age-related increases in depressive symptoms, age-related

decreases in both negative affect and negative social exchanges) compared to the TI. We also found evidence that mode effects for reports of negative affect become more pronounced over time such that as participants age, they show increasingly divergent responses between modes. For reports of positive psychosocial adjustment (i.e., positive affect, positive social exchanges), however, mode effects were not moderated by baseline age, nor did they change over time. These results suggest that age-moderated mode effects may be specific to reports of negative psychosocial adjustment.

Although we assumed that more favorable responding in the IPI compared to the TI is explained by differences in the level of direct contact with the interviewer, other explanations are possible. For example, the IPI were lengthier and more in-depth than the TI, which may have required different levels of involvement and cognitive resources by participants and thereby contributed to different reports across modes. Thus, although the level of direct contact with interviewers is a highly plausible and theoretically grounded mechanism for the consistent pattern of results in the hypothesized directions we found in the LLSSE, we cannot make definitive conclusions about the reasons for mode differences on the basis of these data alone. In Study 2, in contrast to Study 1, participants completed measures of psychosocial functioning with the same assessment method (i.e., self-administered questionnaires) but different levels of direct interviewer contact when returning the questionnaires (i.e., via mail with no direct interviewer contact vs. collected by interviewers during an in-person visit). In Study 2, then, replicating the same pattern of results from Study 1 would bolster support for the interpretation that level of direct interviewer contact may influence participant responses.

Study 2

Study 2 builds on Study 1 by evaluating the same hypotheses with a wider participant age-range (38–86 years old) and longer longitudinal follow-up period (i.e., up to 18 years) examining measures of both negative (i.e., depressive symptoms) and positive psychosocial adjustment (i.e., positive affect). We hypothesized that participants would report fewer depressive symptoms and more positive affect on the data collection method with greater direct contact with interviewers. Following from Study 1, we predicted this effect would be more pronounced with older age, but only for reports of depressive symptoms.

Method

Participants

Participants were from the longitudinal Swedish Adoption/Twin Study of Aging (SATSA; Finkel & Pedersen, 2004; Pedersen et al., 1991). SATSA included same-sex twins from the Swedish Twin Registry who were separated at a young age and reared apart and a sample of same-sex twins who were reared together. Data were collected by mailed questionnaires (Q) and, for a subset of complete pairs aged 50 and older, also by questionnaires collected during an in-person visit by an interviewer (IPTQ; see Procedure). The final sample of 749 participants represents those who took part in the first mailed questionnaire assessing depressive symptoms in 1987 and at least one of four subsequent in-person visits. Participants averaged 60.13 years old (*SD* = 10.48; range = 38–86 years) in 1987 and

included more females than males (59.4% vs. 40.6%). Educational level represented that of Sweden for this cohort, with the majority reporting having completed elementary school level (59%), and the rest having completed ordinary high school (28.3%), gymnasium (6.5%), or university or higher (6.2%)[3].

Procedure

Mailed Questionnaires (Q)—All SATSA participants who had participated in an initial questionnaire assessment in 1984 (Q1; not included in the current paper apart from demographic variables) were mailed a paper-and-pencil questionnaire packet in 1987 (Q2) that they returned in pre-paid envelopes. Q2, but not Q1, included the Center for Epidemiological Studies – Depression Scale (CES-D). Participants completed three additional Q assessments of depressive symptoms in 1990 (Q3), 1993 (Q4), and 2003 (Q5). For these four waves of Q data collection, participants had no direct contact with interviewers; they received questionnaires in the mail and mailed the questionnaires back to the research team.

In-Person Questionnaires (IPTQ)—In-person visits were scheduled between waves of mailed questionnaires. Two weeks prior to the home visit, participants were mailed the identical questionnaire as the Q assessments, including the same depressive symptoms measure. Participants were notified that they should complete the questionnaires prior to the interviewer's in-home visit, who would collect and look over the questionnaire packets to ensure they were completed. The four in-person questionnaire collections (termed "IPTQ" in SATSA) used in the current study occurred: A) between 1989–1991 (IPTQ2), B) between 1992–1994 (IPTQ3), C) between 1999–2001 (IPTQ5), and D) between 2002–2004 (IPTQ6). IPTQ4 was not available (other than a telephone cognitive screening interview). Thus, participants completed up to eight waves of data collection for the current study (four Q assessments and four IPTQ assessments), generally alternating between Q and IPTQ across waves over approximately 18 years.

Measures

All measures were translated into Swedish from English by a bilingual translator and then back-translated into English by a different translator. The translators discussed and resolved any discrepancies between the original and back-translated measures to ensure the original meaning of the measures were retained.

Demographics—Participants' birthdate, gender, and self-reported highest level of education at were available from the Q1 assessment in 1984. Baseline age was based on age at the Q2 assessment in the current study.

[3]Participants who completed at least one in-person questionnaire (IPTQ) in the SATSA (and therefore included in the current analysis) were significantly younger ($M(SD)_{\text{current sample}}$ = 60.13(10.48) vs. $M(SD)_{\text{excluded participants}}$ = 70.06(8.92); $t(155.69) = -10.53$, $p < .001$, unequal variances) and reported fewer health conditions ($M(SD)_{\text{current sample}}$ = 2.56(1.98) vs. $M(SD)_{\text{excluded participants}}$ = 3.15(2.37); $t(103.66) = -2.25$, $p = .027$, unequal variances) and depressive symptoms ($M(SD)_{\text{current sample}}$ = 6.44(6.69) vs. $M(SD)_{\text{excluded participants}}$ = 7.86(7.87); $t(1506.63) = -3.80$, $p < .001$, unequal variances) at the baseline assessment compared to participants who did not complete any IPTQ assessments. Participants in the current sample, however, did not differ from the excluded participants on their education level, $t(819) = .05$, $p = .96$.

Health conditions—The total number of health conditions was derived by summing health problems from 13 different subscales assessing cardiovascular, respiratory, musculoskeletal, allergies, skin, central nervous system, eye, metabolic type, gastrointestinal tract, urologic, cancers, ear, and reproductive organs disorders. Participants reported an average of 2.56 health conditions (*SD* = 1.98, range 0–11) at the baseline Q2 assessment.

Depressive symptoms—Depressive symptoms were assessed using the Center for Epidemiological Studies – Depression Scale (CES-D; Radloff, 1977). We omitted the four positive affect/well-being items from the original version (e.g., "I was happy") in order to have a scale comprised of negative symptoms. Using a four-point scale ranging from 0 (*never or almost never*) to 3 (*always or almost always*), participants rated how often each of the 16 statements was true for them in the last week. Scores at each wave were derived by summing all items for participants who were missing no more than one item on the scale.

Positive Affect—Positive affect was based on the positive affect subscale of the CES-D (Radloff, 1977), comprising 4 items excluded from the core depressive symptoms measure. Participants rated how often the items (related to happiness, enjoying life, being hopeful about the future, and feeling worth as much as others) were true for them on a 4-point scale from 0 (*never or almost never*) to 3 (*always or almost always*). Positive affect scores were based on the sum of the items.

Results

Descriptive Statistics and Preliminary Analyses

Tables 1a and 1b show the descriptive statistics and internal consistencies (as assessed by Cronbach's alpha) for the depressive symptoms and positive affect measures, respectively, at each wave in the study. The Cronbach's alphas were substantial and comparable across waves and data collection modes. Preliminary analyses examining associations between demographic and health variables with depressive symptoms at the baseline Q2 assessment revealed that depressive symptoms were related to age ($r = .18$, $p < .001$), female gender ($r = .11$, $p < .01$), lower educational attainment ($r = -0.10$, $p = .01$), and more health conditions ($r = .26$, $p < .001$). Accordingly, gender (effect coded: −1 = *male*, 1 = *female*), educational level (grand-mean centered), and baseline health conditions (grand-mean centered) were included as covariates in all analyses.

Overview of Data Analysis

Multi-level models were used to account for the hierarchical data structure (i.e., multiple observations nested within individuals nested within twin pairs), with unstructured covariance matrices and the full maximum likelihood method. At Level 1, we examined intra-individual (within-person) variability in depressive symptoms and positive affect across nearly 18 years and up to 8 waves of data collection, and investigated how overall level and change in psychosocial functioning differs by data collection mode (Q vs. IPTQ). Time was calculated at each wave as the number of years since the first assessment (Q2) such that time at the first assessment was zero. Level 2 accounted for the nesting of individuals within twin pairs. Level 3 analyses examined inter-pair (between-pair)

variability (e.g., baseline age differences) in the Level 1 outcome (depressive symptoms, positive affect) by mode. We initially specified both time and mode as random slopes in the model, but the random effect for mode did not reach statistical significance ($p > .05$), so only time was retained as a random effect in the following analyses.

Mode Effects in Self-Reports of Depressive Symptoms and Positive Affect

We hypothesized that participants would provide more favorable responses (i.e., fewer depressive symptoms, greater positive affect) on the mode with more direct contact with the interviewers (i.e., questionnaires collected in-person (IPTQ)) compared to the mode with less direct contact (i.e., questionnaires returned by mail (Q)). In separate models adjusting for baseline age, time, and the covariates (gender, education level, and health conditions), we found a significant main effect of mode (0 = Q, 1 = IPTQ) for reports of depressive symptoms, $B = -1.42$, $SE = 0.15$, $p < .0001$ (accounting for 3% of the explainable within-person variance), as well as positive affect, $B = 0.21$, $SE = 0.07$, $p < .01$ (accounting for about 0.3% of the explainable within-person variance[2]). As predicted, participants reported fewer depressive symptoms and more positive affect in the IPTQ than in the Q.

Tests of Age-Related Moderation of Mode Effects in Reports of Depressive Symptoms

Next, we tested for cross-sectional age moderation whereby mode effects for reports of depressive symptoms may be greater for individuals who are older at the start of the study (i.e., baseline age). We tested this hypothesis using a multi-level model including a cross-product interaction between baseline age (Level 2 variable centered at the sample mean) and data collection mode (Level 1 variable), all lower-ordered main effects (age, mode), and the covariates described previously. The interaction between linear baseline age and data collection mode was not significant, $B = -0.02$, $SE = 0.01$, $p = .11$. Given the relatively large age range of this study sample, however, we examined whether this difference was significant among the oldest adults (as in Study 1) by testing a quadratic age effect. The interaction between quadratic baseline age and mode was significant, $B = -0.003$, $SE = 0.001$, $p = .035$. For IPTQ, reports of depressive symptoms increased linearly across age; for Q, reports of depressive symptoms increased curvilinearly and more steeply with age (see Figure 2A). Thus, the mode effects for reports of depressive symptoms are greatest among the oldest adults.

Additionally, we tested for longitudinal changes to determine whether mode effects become more pronounced over time. Building on the previous model, we included an interaction term between time (in years) and mode, as well as all lower-ordered terms and covariates, predicting depressive symptoms. The Time × Mode interaction was significant, revealing that, as hypothesized, mode effects were enhanced over time (see Table 3 model for depressive symptoms). Specifically, the interaction indicated that although participants generally reported increases in depressive symptoms over time, the mode with greater direct contact with interviewers (IPTQ) produced more favorable estimates of change in depressive symptoms (i.e., less steep increases) compared to the Q which entails less direct contact (see Figure 2B). Exploratory three-way interactions between age (linear and quadratic), time, and mode were not significant, p's > .21.

Tests of Age-Related Moderation of Mode Effects in Reports of Positive Affect

Using the same approach as above, we tested whether cross-sectional baseline age moderates mode effects in reports of positive affect. The Age × Mode interaction did not reach statistical significance, $B = -0.003$, $SE = 0.007$, $p = .70$. We also explored whether, as with the model for depressive symptoms, there may be quadratic age interaction effects, but this interaction was also not significant, $B = 0.00$, $SE = 0.001$, $p = .90$. Next, we investigated whether mode effects in reports of positive affect may become more pronounced over time by including a Time × Mode interaction term with the lower-ordered terms and covariates, but this effect was also not significant (see Table 3 model for positive affect for full model included for illustrative purposes). Thus, similar to Study 1, we did not find evidence that mode effects in reports of positive affect vary with baseline age, nor do they change over time.

Discussion

Replicating the results of Study 1, we found a main effect of mode whereby participants reported fewer depressive symptoms and more positive affect when in greater direct contact with an interviewer. Consistent with our hypothesis, we also found mode effects for reports of depressive symptoms, but not positive affect, were moderated by age and time. Older age at baseline was associated with larger discrepancies in reports of depressive symptoms between the Q and the IPTQ, such that older adults reported more depressive symptoms in the Q compared to the IPTQ, relative to younger individuals. Moreover, participants reported smaller increases over time in depressive symptoms in the mode with greater direct contact with the interviewer (IPTQ) relative to the mode with less direct contact (Q).

Study 2 reproduced the pattern of results from Study 1 even though different sets of modes were used (i.e., Study 1: telephone interviews vs. in-person interviews; Study 2: mailed questionnaires vs. questionnaires collected in-person). In Study 1, the modes differed on more than simply the level of direct contact with interviews (e.g., duration of the interview), which left open other explanations for different reports across the modes (e.g., cognitive demands). In Study 2, however, participants reported their psychosocial functioning via the same methods (i.e., questionnaires), so the cognitive demands were equivalent for both modes. Thus, participants' self-reporting *abilities* are unlikely to explain differences between the modes. Instead, the main difference between the modes was in the level of direct interviewer contact when returning the questionnaires. Given that Study 2 participants were instructed that the interviewer would look over the questionnaire for completeness on one mode but not the other, simply knowing one's answers may be evaluated by an interviewer with whom one is interacting appeared to have a noticeable effect on participants' *willingness* to disclose their psychosocial functioning. These method differences may also explain why the sizes of the main effects for mode were larger in Study 1 compared to Study 2. Together, Study 2 replicated the pattern of results from Study 1, supporting our interpretation that the level of direct interviewer contact influences participants' reports of psychosocial functioning.

General Discussion

In the current set of studies, we found that data collection modes in which participants had greater direct contact with interviewers were associated with reporting less negative [i.e., negative affect, negative social exchanges (Study 1), depressive symptoms (Studies 1 and 2)] and more positive psychosocial adjustment [positive social exchanges (Study 1), positive affect (Studies 1 and 2)]. We also found that in both studies, mode effects in reports of negative, but not positive, psychosocial adjustment were moderated by age. Specifically, older individuals reported less negative psychosocial functioning in modes with greater (vs. less) direct contact with interviewers, relative to younger individuals. Furthermore, we found evidence for longitudinal increases in mode effects for negative, but not positive, psychosocial adjustment (negative affect in Study 1; depressive symptoms in Study 2) across as few as 2 years (Study 1) to 18 years (Study 2). These results indicated that over time, individuals were more sensitive to mode and showed increasingly favorable reporting of negative psychosocial functioning on modes with more direct contact with interviewers compared to modes with less direct contact. Importantly, model estimates derived from modes with greater direct interviewer contact tended to exhibit a pattern of attenuated increases in negative psychosocial functioning over the years compared to modes with less direct contact (discussed in greater detail in the Study Implications section below).

A major contribution of the current study is that we investigated these questions with two different samples, which included participants from different countries and utilized diverse data collection modes, and yet, despite these variations, the pattern of results across the studies was remarkably consistent. Additionally, we longitudinally tracked participants' self-reports of psychosocial functioning across different data collection modes to examine how within-person mode effects may change across the course of years. Accordingly, we were able to examine how potential cohort effects (whereby older adults may have been socialized to refrain from sharing socioemotional problems) and maturational development may uniquely contribute to mode effects. We found that adjusting for the effect of the other, both cohort effects and maturational development independently contributed to mode effects in reports of negative psychosocial development.

In the current studies, age differences and longitudinal change in mode effects were found only for the negative psychosocial measures. As described previously, socioemotional goals become more salient with age (e.g., Carstensen et al., 1999; Riediger et al., 2009) and negative stimuli are often more detrimental to well-being than are positive ones (e.g., Rook, 1984). Thus, reporting on negative psychosocial functioning when in greater direct contact with interviewers may be especially threatening for older adults, and particularly for the oldest-old, for whom losses tend to outweigh gains (e.g., Baltes & Smith, 2003). Older individuals may therefore benefit emotionally by downplaying important negative aspects of the self. This interpretation is consistent with the idea that avoiding negative stimuli may be particularly adaptive for older adults' emotional well-being (e.g., Isaacowitz, Toner, & Neupert, 2009; see reviews by Blanchard-Fields, 2007; Charles, 2010; Charles & Luong, 2013).

With respect to the negative psychosocial measures, we found longitudinal change in mode effects for reports of negative affect in Study 1, but not for depressive symptoms, as we did in Study 2 (negative affect was not assessed in Study 2). One possible explanation for the null finding in Study 1 is that the two studies differed in the duration of their longitudinal investigations. We speculate that because Study 1 only spanned two years, it may have been an insufficient amount of time to detect changes over time in mode effects for depressive symptoms whereas mode effects for reports of negative affect may follow a shorter time course. Additionally, the two measures capture different aspects of psychosocial adjustment; whereas the negative affect items assess mood (e.g., feelings of worry), the measure of depressive symptoms represents substantially poorer psychosocial functioning and somatic problems (e.g., restless sleep and difficulties with concentration and motivation). Thus, while mood states may be more transient, depressive symptoms may show more gradual changes over time. Future studies should examine whether particular measures of psychosocial functioning may show larger longitudinal mode effects and the underlying mechanisms for such changes over time.

Limitations and Future Directions

Some limitations should be noted. First, although our findings were based on rich within-person mixed-method longitudinal studies, interpretations of causal mechanisms should be tempered given the correlational nature of the data. Experimental manipulations may elucidate potential causal pathways and determine whether level of direct interviewer contact explains different reports by mode. In addition, we interpreted longitudinal change in mode effects as reflective of maturational development. It is possible, however, that these findings partially reflect time-in-study effects related to the participant-interviewer relationship dynamics over time. Participants may feel emotionally close to interviewers that they interact with over multiple testing occasions. In these contexts, participants may be especially motivated to provide favorable psychosocial reports with interviewers that they have an ongoing professional relationship with to maintain positive regard with them or to avoid disappointing them. Moreover, participants may genuinely look forward to, and enjoy, socializing with their interviewers. Consequently, participants may provide more favorable psychosocial reports on modes with greater direct contact with interviewers because these visits have a direct and positive impact on the participants' mood. Future studies might delve into participants' and interviewers' evaluations of the interview and their interview partners to test these possibilities.

Additionally, because we were interested in testing longitudinal change in mode effects, participants who did not complete at least one assessment of each mode were excluded from the current analyses. The participants included in the current analyses tended to be younger and reported fewer health conditions and depressive symptoms at the baseline assessments compared to excluded participants (see Footnotes). Thus, another limitation of the current study, as is true of most longitudinal studies, is that there may be selective attrition whereby less healthy individuals may be more likely to drop out of the study. One primary aim of the current investigation, however, was to examine longitudinal change in mode effects, which required the inclusion of participants who completed multiple assessments. Among the younger and relatively healthier participants included in the current investigation, we found

age differences and longitudinal change in mode effects, so it is possible that we actually *underestimated* the mode effects for less healthy older adults, given that the latter individuals may be especially motivated to downplay symptoms of negative psychosocial functioning. This is an empirical question which requires further investigation.

Study Implications

Despite these limitations, the current study offers important insights on how age differences in mode effects may contribute to different age-related patterns of psychosocial functioning. This methodological phenomenon has significant theoretical and conceptual implications for the socioemotional development literature. As discussed in the introduction, there have been some inconsistent findings across studies as to whether psychosocial functioning improves with age. Although mode effects may provide an explanation for some of these discrepancies, much of the socioemotional aging literature has ignored mode effects (but see Soubelet & Salthouse, 2011 for related research). The current study shows that different age-related trajectories of socioemotional adjustment can be obtained from the same participants when assessed via different modes: in the LLSSE, for example, we found no age differences in depressive symptoms in in-person interviews but age-related increases in depressive symptoms in telephone interviews. These results have profound implications because these effects were found within-persons. Given that researchers often make comparisons across different modes and between different samples of age-heterogeneous participants, the inconsistencies in patterns of results may be even more problematic than estimated in the current study.

Our results underscore how mode effects may contribute to divergent patterns of results in the literature and how essential it is for researchers to account for such effects when interpreting findings and formulating theories in the literature. Without doing so, it is possible for researchers using different modes to come up with fundamentally different conclusions about age-related changes in psychosocial functioning. We therefore recommend that instead of aggregating findings across multiple studies for literature reviews and meta-analyses, researchers must consider how the size and direction of such age effects may vary across modes. To date, relatively few reviews have taken the mode of data collection into account, which may contribute to obfuscation of the pattern of age differences in psychosocial adjustment.

Age differences in mode effects also have far-reaching implications for current methodological practices, including how to design future longitudinal investigations, interpret discrepant results in the literature, and analyze mixed-method datasets. For instance, researchers have advocated the use of mixed-method study designs, whereby multiple data collection modes are used within the same study (e.g., Voogt & Saris, 2005; see review by De Leeuw, 2005). Importantly, these recommendations have often been based on cross-sectional studies for which psychosocial development was not of primary interest (but see Schräpler et al., 2010 for an example of a longitudinal validation study). Additionally, some studies confound mode and age effects, such as when younger individuals are administered questionnaires or web surveys whereas older adults are administered telephone interviews (see related discussion by Bowling, 2005). Such practices

may have been based, in part, on assumptions that mode effects are negligible, only applicable to sensitive questions, and/or equal across age groups. For example, Kozma and Stones (1988) stated that "there may be a general tendency by people to overrate their psychological well-being, but as long as this tendency is true for all sub-groups, it is not a major threat to the validity of measures of well-being" (p. 13). Unfortunately, many of these assumptions have often been untested and taken as a given. The current investigation demonstrates that these assumptions do not hold up given that the same individuals will report more favorable psychosocial functioning with modes involving greater direct contact with interviewers – an effect that was stronger with age and time.

Thus, researchers' study design choices may inadvertently influence participant responses and inflate (or deflate) estimates of age-related change in psychosocial adjustment. Life-span developmental researchers should therefore be aware of how their choices in data collection methods may interact with participant aging in ways that influence what participants may be willing to disclose and ultimately, the magnitude and direction of age-related trajectories that may be derived from such data. In scientific reports, we recommend that researchers thoroughly describe participants' level of direct interviewer contact throughout the data collection procedure, given that in the SATSA, even subtle differences in anticipated levels of interviewer contact *after* completing questionnaires (i.e., returning questionnaires via mail with no direct contact vs. returning questionnaires in-person) substantially shaped participants' reports of psychosocial functioning over time, leading to different age-related trajectories in depressive symptoms within the same study. Related to our previous recommendation, sharing this information will allow researchers to conduct literature reviews and meta-analyses of age-related changes to be able to account for, and estimate the influence of, mode effects on the state of the literature. Finally, our results should caution researchers against aggregating data from longitudinal assessments that were collected via different modes, even (or especially) with the same participants, without accounting for mode effects.

Conclusion

The current study provides converging evidence from two longitudinal investigations that data collected from the same individuals using different modes may yield different reports of psychosocial functioning, depending on the level of direct contact with interviewers. We also examined for whom (i.e., older adults) and under which contexts (e.g., longitudinally over time; negative psychosocial measures) these mode effects were most pronounced. The current findings draw attention to the dangers of interpreting data collected via different modes, particularly in longitudinal investigations, without adjusting for mode effects. Researchers are therefore urged to consider the methodological and theoretical implications of how their choices in study designs and data collection methods may influence participant reports of negative psychosocial functioning differentially across age groups and over time.

Acknowledgments

The Later Life Study of Social Exchanges was supported by the National Institute on Aging at the National Institutes of Health (AG14130). The Swedish Twin/Adoption Study of Aging was supported by the National

Institute on Aging (AG04563, AG10175), the MacArthurFoundation Research Network on Successful Aging, and the Swedish Council for Social Research (97:0147:1B).

We would like to thank Michaela Riediger for her helpful comments on an earlier version of the manuscript.

References

Baltes PB, Smith J. New frontiers in the future of aging: From successful aging of the young old to the dilemmas of the fourth age. Gerontology. 2003; 49:123–135. [PubMed: 12574672]

Beekman AT, Copeland JR, Prince MJ. Review of community prevalence of depression in later life. British Journal of Psychiatry. 1999; 174:307–311. [PubMed: 10533549]

Bennett KM. “No sissy stuff”: Towards a theory of masculinity and emotional expression in older widowed men. Journal of Aging Studies. 2007; 21:347–356.

Birditt KS, Fingerman KL, Almeida DM. Age differences in exposure and reactions to interpersonal tensions: A daily diary study. Psychology and Aging. 2005; 20:330–340. [PubMed: 16029096]

Blanchard-Fields F. Everyday problem solving and emotion: An adult developmental perspective. Current Directions in Psychological Science. 2007; 16:26–31.

Bowling A. Mode of questionnaire administration can have serious effects on data quality. Journal of Public Health. 2005; 27:281–291. [PubMed: 15870099]

Carp FM, Carp A. It may not be the answer, it may be the question. Research on Aging. 1981; 3:85–100.

Carstensen LL, Cone JD. Social desirability and the measurement of psychological well-being in elderly persons. Journal of Gerontology. 1983; 38:713–715. [PubMed: 6630907]

Carstensen LL, Isaacowitz DM, Charles ST. Taking time seriously: A theory of socioemotional selectivity. American Psychologist. 1999; 54:165–181. [PubMed: 10199217]

Carstensen LL, Turan B, Scheibe S, Ram N, Ersner-Hershfield H, Samanez-Larkin G, et al. Emotional experience improves with age: Evidence based on over 10 years of experience sampling. Psychology and Aging. 2011; 26:21–33. [PubMed: 20973600]

Charles ST. Strength and vulnerability integration: A model of emotional well-being across adulthood. Psychological Bulletin. 2010; 136:1068–1091. [PubMed: 21038939]

Charles ST, Luong G. Emotional experience across adulthood: The theoretical model of strength and vulnerability integration. Current Directions in Psychological Science. 2013; 22:443–448.

Charles ST, Reynolds CA, Gatz M. Age-related differences and change in positive and negative affect over 23 years. Journal of Personality and Social Psychology. 2001; 80:136–151. [PubMed: 11195886]

Christensen K, Doblhammer G, Rau R, Vaupel JW. Ageing populations: The challenges ahead. Lancet. 2009; 374:1196–1208. doi: http://dx.doi.org/10.1016/S0140-6736(09)61460-4. [PubMed: 19801098]

De Leeuw ED. To mix or not to mix data collection modes in surveys. Journal of Official Statistics. 2005; 21:233–255.

Diener E, Emmons RA. The independence of positive and negative affect. Journal of Personality and Social Psychology. 1984; 47:1105–1117. [PubMed: 6520704]

Dijkstra W, Smit JJ, Comijs HC. Using social desirability scales in research among the elderly. Quality & Quantity. 2001; 35:107–115.

Eaton WW, Kessler LG. Rates of symptoms of depression in a national sample. American Journal of Epidemiology. 1981; 114:528–538. [PubMed: 7304583]

Evans M, Kessler D, Lewis G, Peters TJ, Sharp D. Assessing mental health in primary care research using standardized scales: Can it be carried out over the telephone? Psychological Medicine. 2004; 34:157–162. [PubMed: 14971636]

Finkel D, Pedersen NL. Processing speed and longitudinal trajectories of change for cognitive abilities: The Swedish Adoption/Twin Study of Aging. Aging, Neuropsychology, and Cognition. 2004; 11:325–345.

Gatz M, Hurwicz M-L. Are old people more depressed? Cross-sectional data on center for epidemiological studies depression scale factors. Psychology and Aging. 1990; 5:284–290. [PubMed: 2378694]

Gmel G. The effect of mode of data collection and of non-response on reported alcohol consumption: A split-sample study in Switzerland. Addiction. 2000; 95:123–134. [PubMed: 10723837]

Henderson AS, Jorm AF, Korten AE, Jacomb P, Christensen H, Rodgers B. Symptoms of depression and anxiety during adult life: Evidence for a decline in prevalence with age. Psychological Medicine. 1998; 28:1321–1328. [PubMed: 9854273]

Herman JB. Mixed-mode data collection: Telephone and personal interviewing. Journal of Applied Psychology. 1977; 62:399–404.

Herzog AR, Rodgers W. Interviewing older adults: Mode comparison using data from a face-to-face survey and a telephone survey. Public Opinion Quarterly. 1988; 52:84–89.

Hoher J, Bach T, Munster A, Bouillon B, Tiling T. Does the mode of data collection change results in a subjective knee score? Self-administration versus interview. American Journal of Sports Medicine. 1997; 25:642–647. [PubMed: 9302469]

Isaacowitz DM, Toner K, Neupert SD. Use of gaze for real-time mood regulation: Effects of age and attentional functioning. Psychology and Aging. 2009; 24:989–994. [PubMed: 20025412]

Jorm AF. Does old age reduce the risk of anxiety and depression? A review of epidemiological studies across the adult life span. Psychological Medicine. 2000; 30:11–22. [PubMed: 10722172]

Kessler RC, Birnbaum H, Bromet E, Hwang I, Sampson N, Shahly V. Age differences in major depression: Results from the National Comorbidity Survey Replication (NCS-R). Psychological Medicine. 2010; 40:225–237. [PubMed: 19531277]

Kozma A, Stones MJ. Social desirability in measures of subjective well-being: Age comparisons. Social Indicators Research. 1988; 20:1–14.

Kreuter F, Presser S, Tourangeau R. Social desirability bias in CATI, IVR, and web surveys: The effects of mode and question sensitivity. Public Opinion Quarterly. 2008; 72:847–865.

Kunzmann U, Richter D, Schmukle SC. Stability and change in affective experience across the adult life span: Analyses with a national sample from Germany. Emotion. 2013 Advance online publication.

Lang FR, John D, Lüdtke O, Schupp J, Wagner GG. Short assessment of the Big Five: Robust across survey methods except telephone interviewing. Behavioral Research. 2011; 43:548–567.

Lansford JE, Sherman AM, Antonucci TC. Satisfaction with social networks: An examination of socioemotional selectivity theory across cohorts. Psychology and Aging. 1998; 13:544–552. [PubMed: 9883455]

Lewinsohn PM, Seeley JR, Roberts RE, Allen NB. Center for Epidemoiologic Studies of Depression Scale (CES-D) as a screening instrument for depression among community-residing older adults. Psychology and Aging. 1997; 12:277–287. [PubMed: 9189988]

Luong G, Charles ST, Fingerman KL. Better with age: Social relationships across adulthood. Journal of Social and Personal Relationships. 2011; 28:9–23. [PubMed: 22389547]

Mroczek DK, Kolarz CM. The effect of age on positive and negative affect: A developmental perspective on happiness. Journal of Personality and Social Psychology. 1998; 75:1333–1349. [PubMed: 9866191]

Murrell SA, Himmelfarb S, Wright K. Prevalence of depression and its correlates in older adults. American Journal of Epidemiology. 1983; 117:173–185. [PubMed: 6829547]

Newsom JT, Rook KS, Nishishiba M, Sorkin DH, Mahan TL. Understanding the relative importance of positive and negative social exchanges: Examining specific domains and appraisals. Journal of Gerontology: Psychological Sciences. 2005; 60B:P304–P312.

Nunnally, JC. Psychometric theory. New York: McGraw-Hill; 1978.

Pedersen NL, McClearn GE, Plomin R, Nesselroade JR, Berg S, DeFaire U. The Swedish Adoption/Twin Study of Aging: An update. Acta Geneticae Medicae et Gemellologiae: Twin Research. 1991; 40:7–20.

Pinquart M. Age differences in perceived positive affect, negative affect, and affect balance in middle and old age. Journal of Happiness Studies. 2001; 2:375–405.

Presser S, Stinson L. Data collection mode and social desirability bias in self- reported religious attendance. American Sociological Review. 1998; 63:137–145.

Radloff LS. The CES-D Scale: A self-report depression scale for research in the general population. Applied Psychological Measurement. 1977; 1:385–401.

Ray JJ. Lie scales and the elderly. Personality and Individual Differences. 1988; 9:417–418.

Richman WL, Kiesler S, Weisband S, Drasgow F. A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. Journal of Applied Psychology. 1999; 84:754–775.

Riediger M, Schmiedek F, Wagner GG, Lindenberger U. Seeking pleasure and seeking pain: Differences in prohedonic and contra-hedonic motivation from adolescence to old age. Psychological Science. 2009; 20:1529–1535. [PubMed: 19891749]

Rook KS. The negative side of social interaction: Impact on psychological well-being. Journal of Personality and Social Psychology. 1984; 46:1097–1108. [PubMed: 6737206]

Santor DA, Coyne JC. Shortening the CES-D to improve its ability to detect cases of depression. Psychological Assessment. 1997; 9:233–243.

Schräpler J-P, Schupp J, Wagner GG. Changing from PAPI to CAPI: Introducing CAPI in a longitudinal study. Journal of Official Statistics. 2010; 26:239–269.

Singer JD. Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. Journal of Educational and Behavioral Statistics. 1998; 24:323–355.

Snowdon J. Is depression more prevalent in old age? Australian and New Zealand Journal of Psychiatry. 2001; 35:782–787. [PubMed: 11990889]

Sorkin DH, Rook KS. Interpersonal control strivings and vulnerability to negative social exchanges in later life. Psychology and Aging. 2004; 19:555–564. [PubMed: 15584782]

Soubelet A, Salthouse TA. Influence of social desirability on age differences in self-reports of mood and personality. Journal of Personality. 2011; 79:741–762. [PubMed: 21682727]

Spinhoven PH, Ormel J, Sloekers PPA, Kempen GIJM, Speckens AEM, Van Hemert AM. A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. Psychological Medicine. 1997; 27:363–370. [PubMed: 9089829]

Stordal E, Bjartveit Krüger M, Dahl NH, Mykletun A, Dahl AA. Depression in relation to age and gender in the general population: The Nord-Trøndelag Health Study (HUNT). Acta Psychiatrica Scandinavica. 2001; 104:210–216. [PubMed: 11531658]

Thomsen DK, Mehlsen MY, Viidik A, Sommerlund B, Zachariae R. Age and gender differences in negative affect - Is there a role for emotion regulation? Personality and Individual Differences. 2005; 38:1935–1946.

Tourangeau R, Smith TW. Asking sensitive questions: The impact of data collection mode, question format, and question context. The Public Opinion Quarterly. 1996; 60:275–304.

Turvey C, Wallace RB, Herzog R. A revised CES-D measure of depressive symptoms and a DSM-based measure of depressive episodes in the elderly. International Psychogeriatrics. 1999; 11:139–148. [PubMed: 11475428]

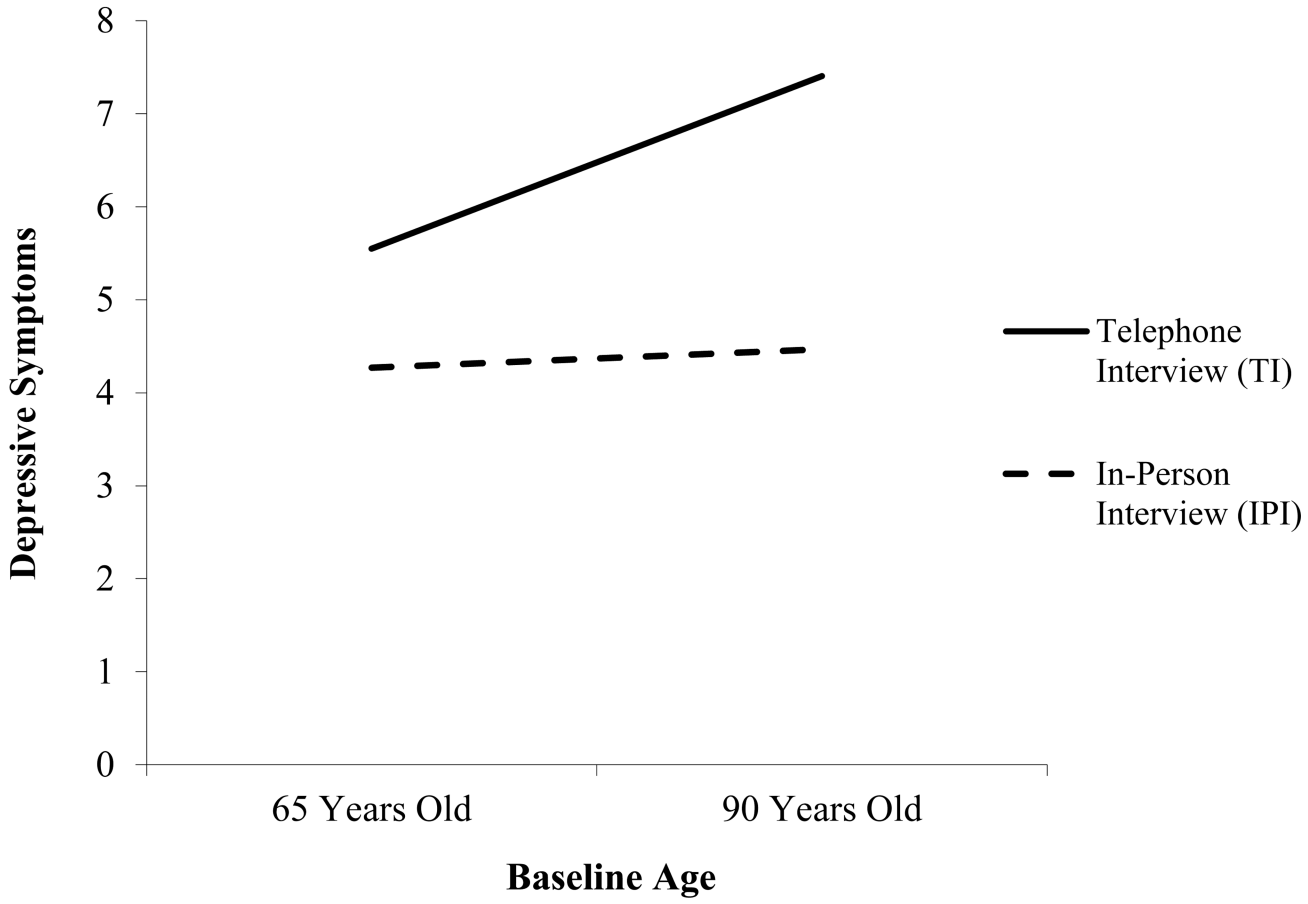
Unützer J. Diagnosis and treatment of older adults with depression in primary care. Biological Psychiatry. 2002; 52:285–292. doi: http://dx.doi.org/10.1016/S0006-3223(0201338-0). [PubMed: 12182933]

Vannieuwenhuyze J, Looseveldt G, Molenberghs G. A method for evaluating mode effects in mixed-mode surveys. Public Opinion Quarterly. 2010; 74:1027–1045.

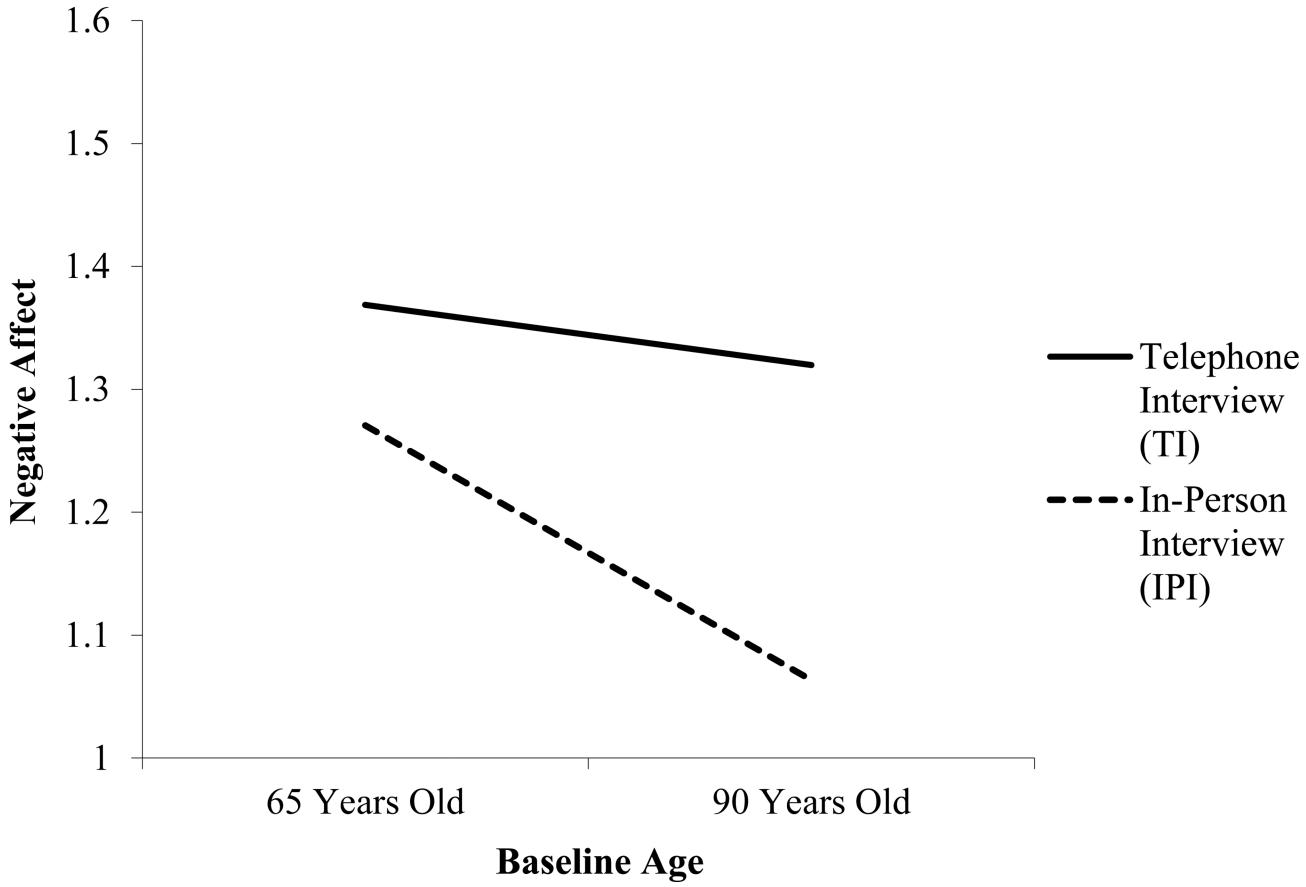
von Hippel W, Henry JD, Matovic D. Aging and social satisfaction: Offsetting positive and negative effects. Psychology and Aging. 2008; 23:435–439. [PubMed: 18573016]

Voogt RJJ, Saris WE. Mixed mode designs: Finding the balance between nonresponse bias and mode effects. Journal of Official Statistics. 2005; 21:367–387.

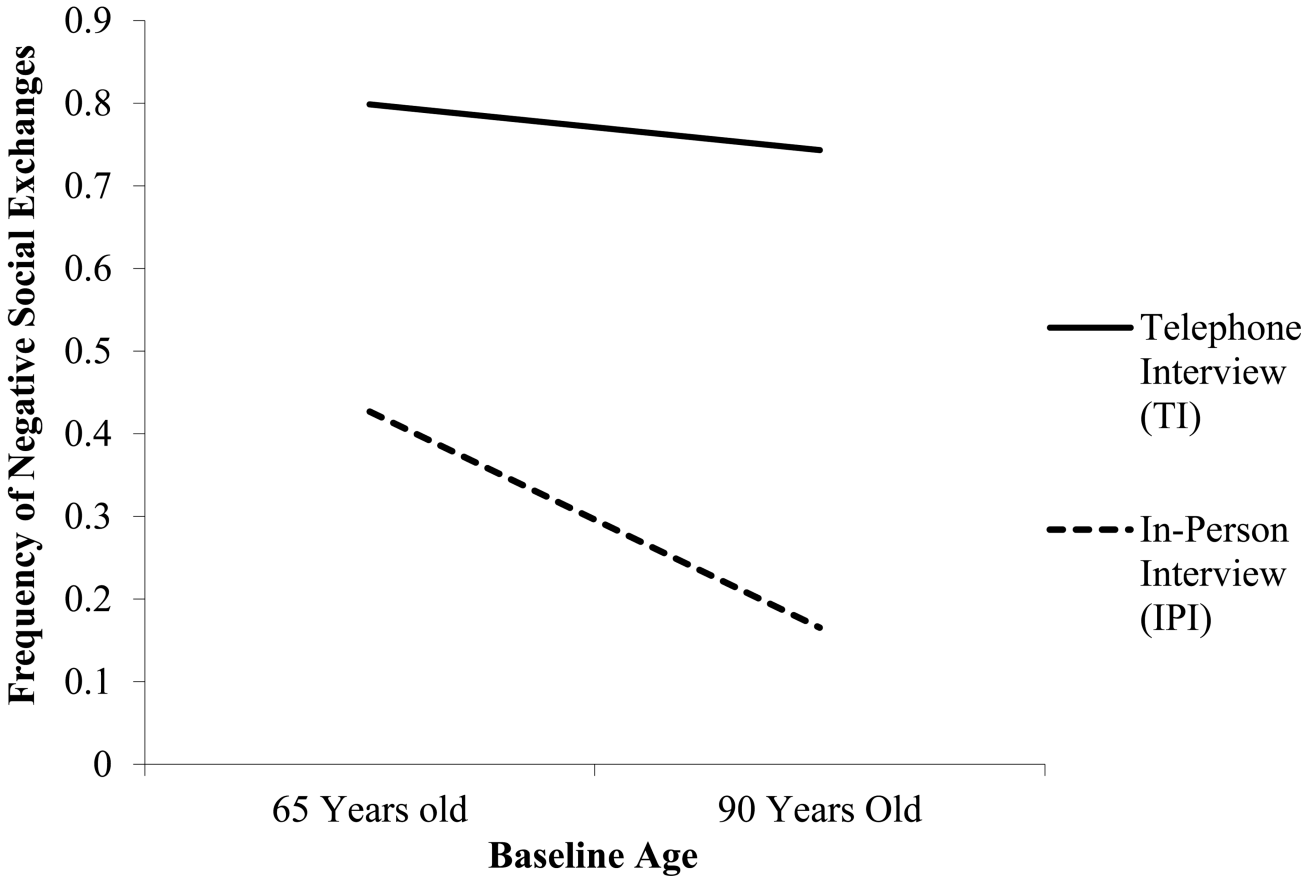
A



B



C



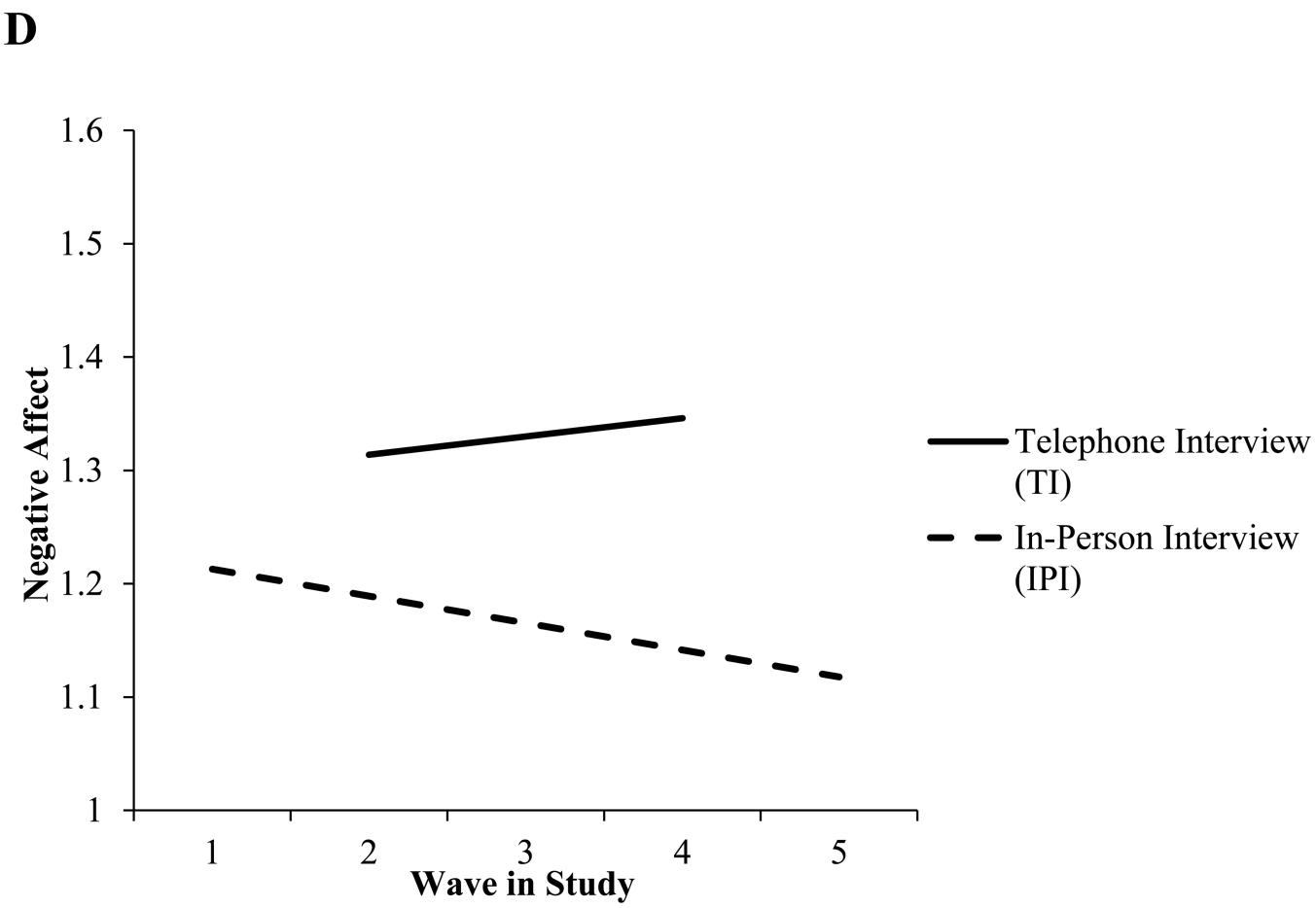


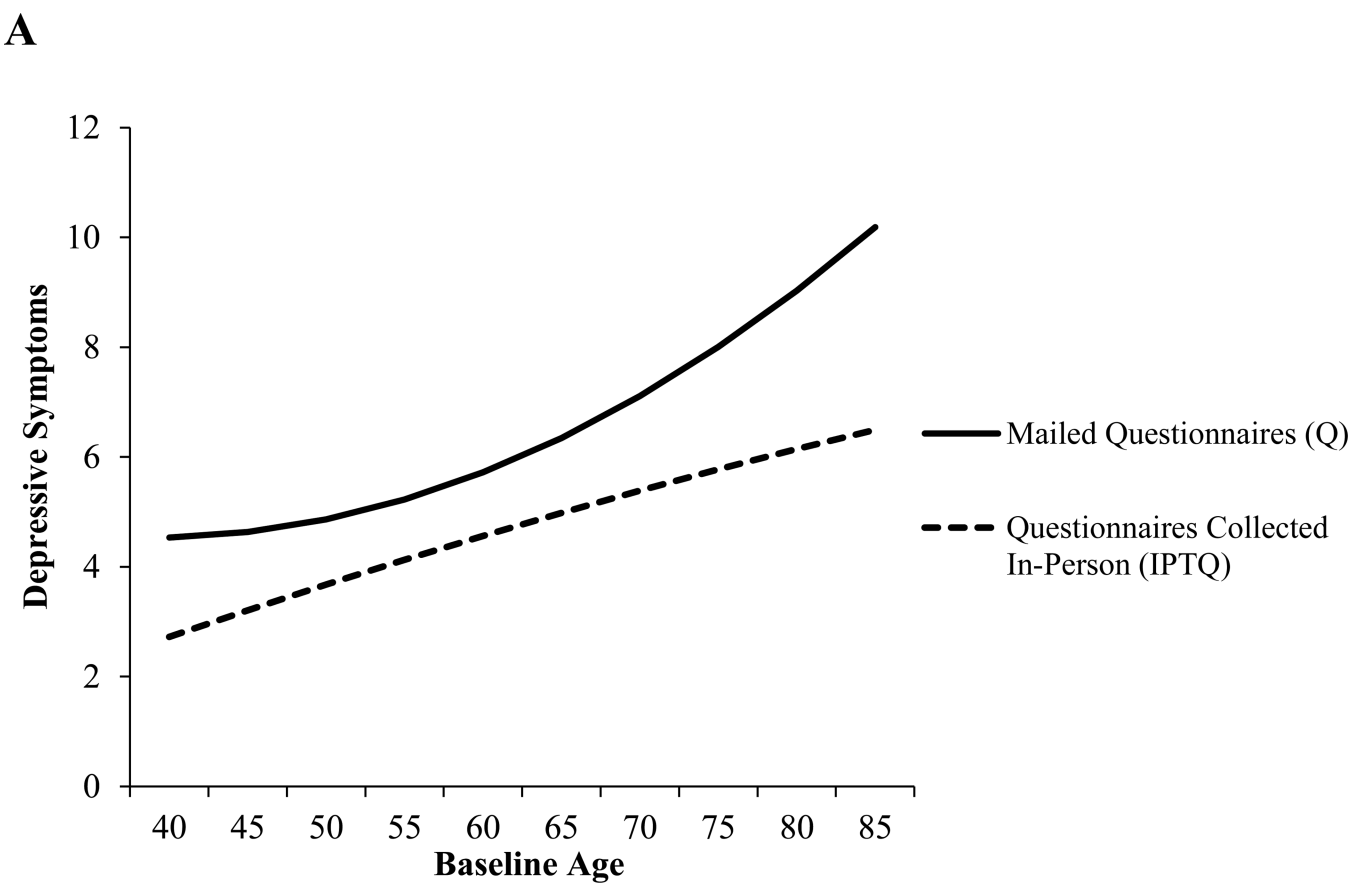
Figure 1.

A. Study 1: Baseline Age by Data Collection Method Interaction Predicting Reports of Depressive Symptoms in the LLSSE

B. Study 1: Baseline Age by Data Collection Method Interaction Predicting Reports of Negative Affect in the LLSSE

C. Study 1: Baseline Age by Data Collection Method Interaction Predicting Reports of Negative Social Exchanges in the LLSSE

D. Study 1: Time by Data Collection Method Interaction Predicting Reports of Negative Affect in the LLSSE



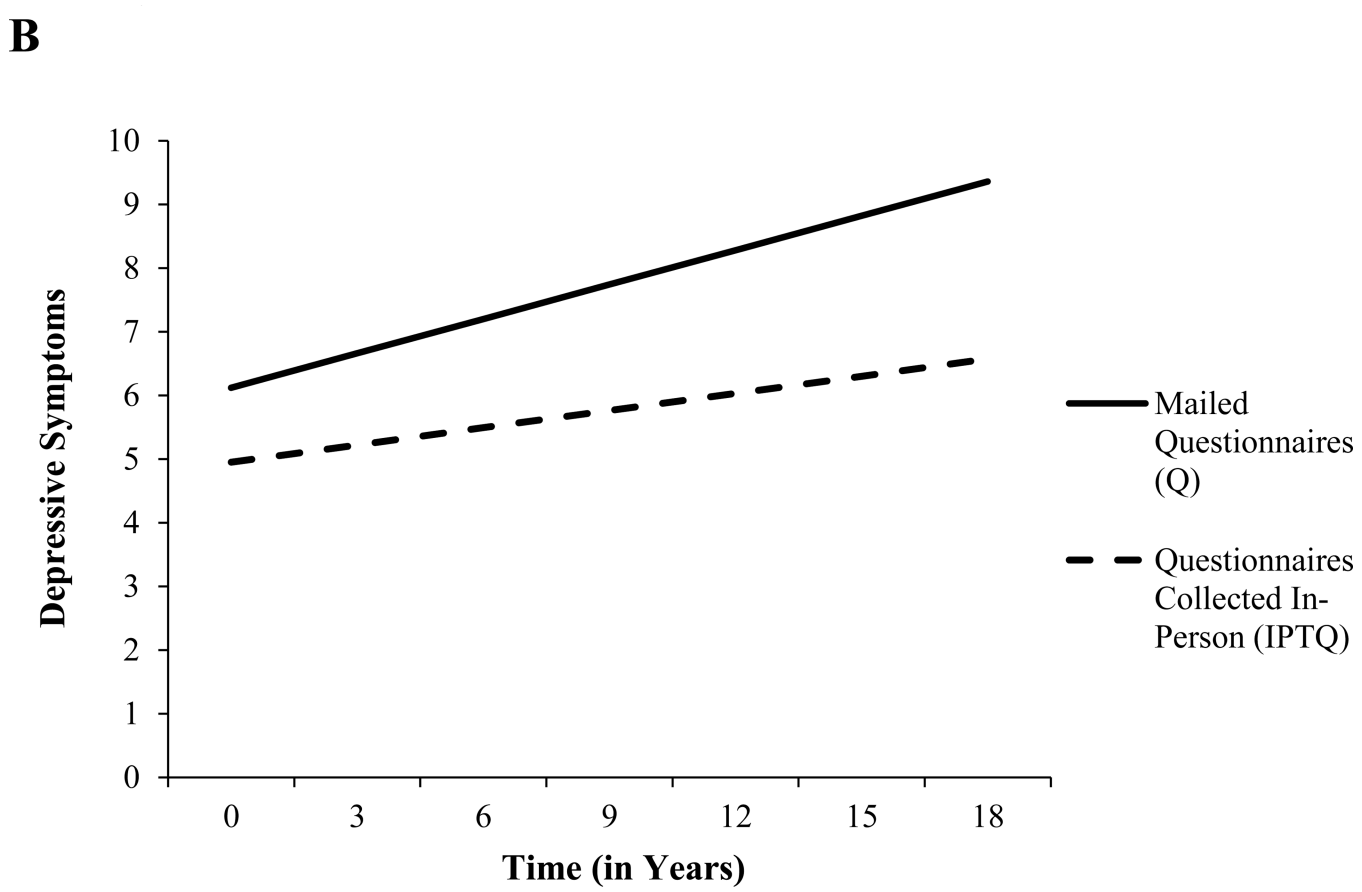


Figure 2.

A. Study 2: Quadratic Baseline Age by Data Collection Mode Interaction Predicting Reports of Depressive Symptoms in the SATSA

B. Study 2: Time by Data Collection Mode Interaction Predicting Reports of Depressive Symptoms in the SATSA

Table 1

a. Descriptive Statistics of Negative Psychosocial Functioning Across Modes and Waves in the LLSSE and SATSA

	LLSSE Depressive Symptoms (9 items)			LLSSE Negative Affect			LLSSE Negative Social Exchanges			SATSA Depressive Symptoms (16 items)		
	Mode	***M*(*SD*)**	**α**	**Mode**	***M*(*SD*)**	**α**	**Mode**	***M*(*SD*)**	**α**	**Mode**	***M*(*SD*)**	**α**
Wave 1	IPI	4.87(4.70)	.82	IPI	1.34(0.76)	.80	IPI	0.42(0.56)	.89	Q	6.44(6.69)	.90
Wave 2	TI	6.94(5.18)	.84	TI	1.43(0.72)	.78	TI	0.81(0.65)	.89	IPTQ	6.68(6.84)	.90
Wave 3	IPI	5.28(4.71)	.83	IPI	1.27(0.72)	.80	IPI	0.44(0.57)	.90	Q	6.86(6.25)	.88
Wave 4	TI	6.96(5.15)	.84	TI	1.46(0.70)	.78	TI	0.80(0.63)	.90	IPTQ	6.11(5.87)	.86
Wave 5	IPI	5.21(4.85)	.83	IPI	1.25(0.71)	.81	IPI	0.43(0.58)	.91	Q	7.73(7.20)	.90
Wave 6	---	---	---	---	---	---	---	---	---	IPTQ	6.36(6.38)	.89
Wave 7	---	---	---	---	---	---	---	---	---	Q	8.27(6.99)	.89
Wave 8	---	---	---	---	---	---	---	---	---	IPTQ	6.57(6.41)	.88

b. Descriptive Statistics of Positive Psychosocial Functioning Across Modes and Waves in the LLSSE and SATSA

	LLSSE Positive Affect (5 items)			LLSSE Positive Social Exchanges			SATSA Positive Affect (4 items)		
	Mode	***M*(*SD*)**	**α**	**Mode**	***M*(*SD*)**	**α**	**Mode**	***M*(*SD*)**	**α**
Wave 1	IPI	3.02(0.66)	.85	IPI	2.36(.089)	.89	Q	8.05(2.73)	.67
Wave 2	TI	2.92(0.71)	.86	TI	2.36(0.69)	.86	IPTQ	8.06(2.77)	.68
Wave 3	IPI	2.89(0.71)	.88	IPI	2.45(0.79)	.88	Q	8.08(2.66)	.67
Wave 4	TI	2.84(0.69)	.87	TI	2.35(0.68)	.87	IPTQ	8.34(2.69)	.65
Wave 5	IPI	2.92(0.68)	.89	IPI	2.46(0.82)	.89	Q	7.98(2.72)	.71
Wave 6	---	---	---	---	---	---	IPTQ	7.86(2.88)	.66
Wave 7	---	---	---	---	---	---	Q	7.82(2.57)	.67
Wave 8	---	---	---	---	---	---	IPTQ	8.09(2.84)	.67

Notes. Mode refers to data collection mode. IPI = in-person interview. TI = telephone interview. Q = mailed questionnaire. IPTQ = questionnaires collected in-person. α = Cronbach's alpha. In the Later Life Study of Social Exchanges (LLSSE), the measure of depressive symptoms was based on a 9-item version of the CES-D (possible score range 0–27). In the Swedish Adoption/Twin Study of Aging (SATSA), the measure of depressive symptoms was based on a 16-item version of the CES-D (possible score range 0–48).

Notes. Mode refers to data collection mode. IPI = in-person interview. TI = telephone interview. Q = mailed questionnaire. IPTQ = questionnaires collected in-person. α = Cronbach's alpha. In the LLSSE, the measure of positive affect was based on a 5-item scale developed by Diener and Emmons (1984) with a possible score range from 0 to 5. In the SATSA, the measure of positive affect was based on 4-item version of the CES-D subscale (possible score range 0 to 12).

Table 2

Study 1: Negative and Positive Psychosocial Functioning Regressed on Age, Time, and Data Collection Mode in the LLSSE

	Negative Psychosocial Functioning						Positive Psychosocial Functioning			
	DV = Depressive Symptoms		DV = Negative Affect		DV = Negative Social Exchanges		DV = Positive Affect		DV = Positive Social Exchanges	
Fixed Effects	***Estimate***	***SE***	***Estimate***	***SE***	***Estimate***	***SE***	***Estimate***	***SE***	***Estimate***	***SE***
Intercept	6.75***	0.23	1.39***	0.03	0.82***	0.03	2.97***	0.03	2.33***	0.04
Gender (1 = *female*)	0.48***	0.14	0.09***	0.02	0.01	0.02	0.01	0.02	0.09***	0.02
Education	−0.16*	0.07	−0.01	0.01	−0.01	0.01	0.00	0.01	−0.02	0.01
Chronic Health Conditions	0.72***	0.09	0.08***	0.01	0.05***	0.01	−0.07***	0.01	0.01	0.01
Mode (1 = *IPI*)	−1.96***	0.22	−0.08**	0.03	−0.39***	0.03	0.03	0.03	0.01	0.03
Time	0.07	0.09	0.02	0.01	−0.01	0.01	−0.05***	0.01	0.00	0.01
Baseline Age	0.07**	0.02	0.00	0.00	0.00	0.00	−0.01*	0.00	0.00	0.00
Baseline Age × Mode	−0.06**	0.02	−0.01*	0.00	−0.01***	0.00	0.00	0.00	0.00	0.00
Time × Mode	0.06	0.09	−0.04*	0.01	0.01	0.01	0.02	0.01	0.03†	0.02
Random Effects										
Level 2 (Between Persons)										
Intercept	15.04***	1.29	0.31***	0.03	0.25***	0.02	0.24***	0.02	0.28***	0.03
Time	0.26***	0.08	0.01***	0.00	0.001†	0.001	0.002†	0.002	0.02***	0.00
Mode	4.14***	0.74	0.05***	0.01	0.11***	0.01	0.03*	0.01	0.12***	0.02
COV (Intercept, Time)	−0.29	0.24	−0.01**	0.01	0.00	0.00	0.003	0.005	−0.03***	0.01
COV (Intercept, Mode)	−4.26***	0.76	−0.02†	0.01	−0.10***	0.01	−0.02†	0.01	0.02	0.02
COV (Time, Mode)	0.13	0.15	0.00	0.00	0.00	0.00	0.00	0.00	−0.01	0.00
Level 1 (Within Persons)										
Residual	9.70***	0.37	0.21***	0.01	0.14***	0.01	0.21***	0.01	0.25***	0.01

Notes. DV = dependent variable. For data collection mode, 0 = telephone interviews (TI) and 1 = in-person interviews (IPI). Gender is effect coded (−1 = male, 1 = female). Time refers to wave in the study. COV refers to the covariance parameter between the specified random effects.

$^{\dagger}p < .10$;

Author Manuscript

Author Manuscript

Author Manuscript

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

Table 3

Study 2: Depressive Symptoms and Positive Affect Regressed on Baseline Age, Time, and Data Collection Mode in the SATSA

	DV = Depressive Symptoms		DV = Positive Affect	
Fixed Effects	***Estimate***	***SE***	***Estimate***	***SE***
Intercept	6.12^{***}	0.32	8.15^{***}	0.13
Gender (1 = *female*)	0.58^{**}	0.21	0.08	0.08
Education	−0.19	0.23	0.26^{**}	0.09
Chronic Health Conditions	0.79^{***}	0.10	-0.18^{***}	0.09
Mode (1 = *IPTQ*)	-1.17^{***}	0.20	0.20^{*}	0.10
Time	0.18^{***}	0.02	-0.02^{*}	0.01
Baseline Age	0.12^{***}	0.02	−0.01	0.01
Baseline Age^2	0.003	0.002	−0.001	0.001
Baseline Age × Mode	-0.04^{**}	0.02	−0.005	0.007
Baseline Age^2 × Mode	-0.003^{*}	0.001	0.000	0.001
Time in Study × Mode	-0.09^{***}	0.03	−0.01	0.01
Random Effects				
Level 3 (Between Twin Pairs)				
Intercept	8.94^{***}	1.98	0.81^{**}	0.31
Time	0.02^{*}	0.01	$0.002^{\dagger}$	0.002
COV(Intercept, Time)	-0.30^{*}	0.12	−0.01	0.02
Level 2 (Within Twin Pairs)				
Intercept	14.75^{***}	1.93	2.44^{***}	0.36
Time	0.05^{***}	0.01	0.002	0.002
COV(Intercept, Time)	-0.32^{*}	0.13	−0.03	0.02
Level 1 (Within Persons)				
Residual	18.04^{***}	0.48	4.26^{***}	0.11

Notes. For data collection mode, 0 = mailed questionnaires (Q) and 1 = questionnaires collected in-person (IPTQ). Gender is effect coded (−1 = male, 1 = female). Time refers to the number of years since the baseline assessment. COV refers to the covariance parameter between the specified random effects.

$^{\dagger}p < .10$;

$^{*}p < .05$;

$^{**}p < .01$;

$^{***}p < .001$.