Article

Longitudinal Changes in Empathy Across the Life Span in Six Samples of Human Development

Social Psychological and Personality Science
1-10

DOI: 10.1177/1948550619849429
journals.sagepub.com/home/spp
SAGE

Jeewon Oh[1], William J. Chopik[1], Sara Konrath[2], and Kevin J. Grimm[3]

Abstract

The development of empathy is a hotly debated topic. Some studies find declines and others an inverse U-shaped pattern in empathy across the life span. Yet other studies find no age-related changes. Most of this research is cross sectional, and the few longitudinal studies have their limitations. The current study addresses these limitations by examining changes in empathy in six longitudinal samples (total $N = 740$, age 13–72). In a preliminary study ($N = 784$), we created and validated a measure of empathy out of the California Adult Q-Sort. The samples were combined for multilevel analyses in a variant of an accelerated longitudinal design. We found that empathy increased across the life span, particularly after age 40, and more recent cohorts were higher in empathy.

Keywords

empathy, life span development, Q-Sorts, personality

Empathy is the ability to perceive and feel others' thoughts and feelings (Davis, 1983; Hoffman, 2008). Researchers have often focused on two components of empathy: the cognitive (perceiving and considering others' perspectives) and emotional components (feeling concern and compassion for others; Bailey, Henry, & Von Hippel, 2008). Although empathy has been assumed to be an enduring individual characteristic (Eisenberg et al., 1999; Grühn, Rebucal, Diehl, Lumley, & Labouvie-Vief, 2008), there is also a debate about whether and how it changes as people age. Unfortunately, many studies have relied on cross-sectional data (O'Brien, Konrath, Grühn, & Hagen, 2013), resulting in some ambiguity about whether empathy changes with age (Sze, Gyurak, Goodkind, & Levenson, 2012) or whether preexisting differences between birth cohorts explain age differences in empathy (Konrath, O'Brien, & Hsing, 2011). The current study examined lifelong changes in empathy from six samples of individuals from ages 13 to 72, who were born over a 46-year period (1923–1969).

Theories of Empathy Development

The cognitive and relational development that occurs during adolescence might lead to changes in empathy (Van der Graaff et al., 2016). Through emerging adulthood, people continue to develop a consideration for others and their responsibilities as independent adults (Arnett, 1998) that may help them better understand others and regulate their interpersonal interactions, ultimately increasing empathy. However, why would one expect empathy to change with age long after this initial development? Developmental researchers suggest that individuals change in ways that are consistent with a "maturation" that occurs across the entirety of young adulthood and through middle age, and these changes reflect a functional adaptation that yields success in work and relationships (Roberts & Mroczek, 2008). Thus, as individuals age, they have a greater motivation to empathize with others to cultivate successful relationships. Further, socioemotional selectivity theory (SST) suggests that as people age and experience certain life events (e.g., retirement), they see time as limited and are motivated to achieve more emotion-related goals, which may be facilitated by empathy (Carstensen, Isaacowitz, & Charles, 1999). For instance, older adults focus more on investing in close relationships that help serve emotional goals (Carstensen et al., 1999; Chopik, 2017), and when experimental studies manipulate relational closeness, empathy increases with enhanced closeness (Zhang, Fung, Stanley, Isaacowitz, & Ho, 2013). Indeed, empathy is related to constructs that are aligned with satisfying emotional goals: Empathy is positively associated with relationship satisfaction (Davis & Oathout, 1987) and

[1] Michigan State University, East Lansing, MI, USA
[2] Lilly Family School of Philanthropy, Indiana University, Indianapolis, IN, USA
[3] Arizona State University, Tempe, AZ, USA

Corresponding Author:
Jeewon Oh, Department of Psychology, Michigan State University, 316 Physics Rd., East Lansing, MI 48824, USA.
Email: ohjeewon@msu.edu

negatively associated with depression and relational conflict (Cramer & Jowett, 2010).

However, aging may introduce declines in cognitive ability that hinder perspective taking (Bailey et al., 2008; Ruffman, Henry, Livingstone, & Phillips, 2008). Inhibitory control becomes increasingly difficult with age, making it harder to inhibit natural and automatic thoughts to consider others' thoughts and feelings (Bailey & Henry, 2008). Therefore, cognitive empathy may decline with age later in life.

Existing Evidence for Age Differences and Age-Related Changes in Empathy

Some short-term longitudinal studies of adolescents find increases in empathy (Allemand, Steiger, & Fend, 2015; Davis & Franzoi, 1991; Van der Graaff et al., 2016). Other studies find that older adults have greater emotional empathy (Sze et al., 2012). Developmental researchers and SST suggest that individuals change in ways that yield success in work and relationships (Carstensen et al., 1999; Roberts & Mroczek, 2008). Thus, people may be more motivated to empathize with others to cultivate successful relationships as they age.

There is also evidence that empathy peaks in midlife before declining in late life (O'Brien et al., 2013). This switch in late life introduces another perspective that empathy might *decrease* with age (Helson, Jones, & Kwan, 2002). Specifically, studies find that emotion recognition from faces generally declines with age (Ruffman et al., 2008). Further, while cognitive empathy does not change with age, emotional empathy is affected by neuroanatomical changes that occur in late life (Chen, Chen, Decety, & Cheng, 2013). In fact, one of the few long-term longitudinal studies of empathy shows that empathy declines from age 21 to 60 using a proprietary measure of empathy (the California Psychological Inventory (CPI); Helson et al., 2002). Thus, neuroanatomical changes in the brain might explain why empathy is negatively associated with age in some studies, particularly in older adulthood (Cabello, Sorrel, Fernandez-Pinto, Extremera, & Fernandez-Berrocal, 2016; O'Brien et al., 2013).

Although some studies suggest increases or decreases in empathy, the preponderance of evidence suggests that empathy shows little or no change across the life span (Beadle, Sheehan, Dahlben, & Gutchess, 2015; Eysenck, Pearson, Easting, & Allsopp, 1985; Grühn et al., 2008; Hannikainen, Machery, & Cushman, 2018). Even among studies that find age differences, these conclusions are tempered when changes are specific to a component of empathy, not found when examining broad age ranges, are only found in some samples even within the same reports or are context dependent (Bailey et al., 2008; Grühn et al., 2008; Helson et al., 2002; Van der Graaff et al., 2016; Wieck & Kunzmann, 2015).

Methodological Limitations of Previous Work

The literature's reliance on cross-sectional data has led to some ambiguity about whether age differences in empathy are due to cohort effects or developmental processes (Grühn et al., 2008; Konrath et al., 2011; Labouvie-Vief, Grühn, & Studer, 2010). Researchers examining generational and cohort changes in empathy also arrive at different conclusions—some suggest that empathy might be higher in recent generations (e.g., Grühn et al., 2008) and others lower in recent generations (Hannikainen et al., 2018; Konrath et al., 2011). Proposed reasons for these increases and decreases range from social networking use, media and technology consumption, and changes in other psychological characteristics and behavior, parenting and family practices, and the broader cultural zeitgeist (e.g., a culture of success, a culture of social consciousness; Chopik, Joshi, & Konrath, 2014; Greenberg & Weber, 2008; Grühn et al., 2008; Howe & Strauss, 2000; Konrath et al., 2011; Rifkin, 2010; Twenge, 2006).

The Current Study

Our existing knowledge for how empathy changes over long intervals beyond adolescence is based primarily on two studies that used the same proprietary measure of empathy (the CPI; Grühn et al., 2008; Helson et al., 2002). Unfortunately, information about the contents of this measure, the items (and coding scheme) are limited and not readily available to researchers, but studies that use other empathy measures (e.g., the Interpersonal Reactivity Index [IRI]; Davis, 1983) occasionally find increases in empathy. The IRI suggests empathy is a multidimensional construct with four subscales (fantasy, perspective taking, empathic concern, and personal distress), but developmental researchers have mostly focused on perspective taking (cognitive empathy) and empathic concern (emotional empathy). Although this measure shows good reliability and validity (Gilet, Mella, Studer, Grühn, & Labouvie-Vief, 2013; Van der Graaff et al., 2016), there are few longitudinal studies and none that follow individuals from adolescence to older adulthood (Davis & Franzoi, 1991). Other measures were constructed based on expert judgments of the construct of empathy but not properly validated using other measures or situated within a larger nomological network (Hogan, 1969). Differences in item content, sample characteristics, or the narrow age range in prior research could explain the discrepant findings regarding age-related changes in empathy. Given that long-term longitudinal studies are rare and existing studies have their own limitations, it is difficult to understand the relationship between age and empathy.

To address these limitations, we measured empathy from ages 13 to 72 using a new measure of empathy derived from the California Adult Q-Sort (CAQ), which was included in six longitudinal studies. This measure was based on observer ratings, rather than self-report, which minimizes concerns about social desirability or a lack of self-insight. We consider the use of observer ratings to be a major strength of the current study. Recent research has highlighted the utility of examining observer reports of psychological characteristics in developmental research. Observers' insights can often add more accurate and predictive information than self-reports of psychological characteristics, particularly if those characteristics are observable

and evaluative (e.g., charming, funny; Vazire & Carlson, 2011; Vazire & Mehl, 2008). Empathy can be considered an evaluative trait with observable behaviors, so observer ratings can provide useful information for how individuals vary in empathy while sidestepping many of the limitations of longitudinal studies (e.g., retest effects).

Method

Participants

Block and Block Longitudinal Study of Cognitive and Ego Development. One hundred seven people (50.5% female) participated in the Block and Block Longitudinal Study of Cognitive and Ego Development, which was initiated in 1968 at the University of California at Berkeley (for full description, see Block & Block, 2006). The sample was recruited from two preschools in Berkeley, CA. CAQs (Block, 2008) were collected at ages 14, 18, and 23. The ethnic composition of the sample was 68.3% Caucasian, 24% African American, 4.8% Asian American, and 2.9% Other ethnicities.

Intergenerational studies. Three hundred fifty-four people (53.7% female) participated in the intergenerational studies, an umbrella study combining participants from the Berkeley Guidance Study ($N = 171$), the Oakland Growth Study ($N = 121$), and the Berkeley Growth Study ($N = 62$). The three longitudinal studies were started in the late 1920s and early 1930s and continued for over 70 years. The Berkeley Guidance and Growth Studies sampled infants born in the Berkley area in 1928–1929. The Oakland Growth Study began in 1932 and sampled fifth and sixth graders (approximate birth year = 1921). Frequency of CAQ assessments was high in each study: Berkeley Guidance (ages 13, 16, 30, 40, and 52), Oakland Growth (ages 13, 16, 40, 50, and 60), and Berkeley Growth (ages 13, 16, 36, and 52). All participants from the Berkeley Growth and Oakland Growth Study were Caucasian; a small percentage of participants from the Guidance Study were African American (3%). The intergenerational studies are considered landmark studies in human development (see Block, 1971; Eichorn, Clausen, Haan, Honzik, & Mussen, 1981; Haan, Millsap, & Hartka, 1986, for details on their history and sampling).

Mills longitudinal study. One hundred and twenty-two female individuals participated in a longitudinal study of the 1958 and 1960 senior classes at Mills College (for a full description, see Helson, 1993). CAQs were collected at ages 21 and 43. The sample was predominantly Caucasian; exact percentages of racial/ethnic groups are unavailable.

Radcliffe College class of 1964. One hundred and sixty-seven female individuals participated in a longitudinal study of members of the 1964 graduating class of Radcliffe College (see Stewart & Vandewater, 1993, for full description). CAQs were collected at ages 43, 53, 62, and 72. All but one woman were European American.

Combined sample. Table 1 presents sample sizes at each assessment wave, study design, and the number and timing of assessment points. The average age of the combined sample ($N = 740$) was 36.65 years ($SD = 18.04$; 68.4% female) across assessment waves. Across all participants and all assessment waves, there were 2,101 observations. We included all available data and did not exclude any participants in the analyses. A sensitivity analysis suggested that given 80% power at $\alpha = .05$, the smallest possible effect we could estimate is Cohen's $f^2 = .005$.

Empathy Measure

At each wave, empathy was assessed using items from the CAQ (Block, 1961, 2008). CAQ is a broad personality measure comprised of 100 descriptive items. These items are sorted using a forced distribution into nine categories (from 1 = *extremely uncharacteristic* to 9 = *extremely characteristic*). Evaluations were completed by study personnel, who included clinicians, graduate students, and psychologists. CAQ for the Block and Block Study was completed primarily by the research investigators and research scientists working on the project; the Radcliffe and Mills sample by trained graduate students and research scientists; and the intergenerational studies by trained clinicians and study personnel. In each case, multiple sorters (often two or three) provided ratings on each participant, and interrater agreement was on average .78 across all samples and waves. The evaluations were based on in-depth interviews and observations conducted during a variety of experimental tasks. CAQs from multiple observers were averaged within each participant at each assessment wave.

Because there was no validated study of empathy measured through the CAQ, we employed the following procedure to identify items that could conceivably measure empathy (Chopik & Edelstein, 2015; Newton & Stewart, 2013): Specifically, we first correlated individual CAQ items with composite measures of empathic concern ($\alpha = .81$), perspective taking ($\alpha = .78$), and an aggregate score that combined empathic concern and perspective taking ("total empathy"; $\alpha = .84$; see Chopik, O'Brien, & Konrath, 2017) from a contemporary measure of empathy (the IRI; Davis, 1983) in an undergraduate student sample ($N = 784$, $M_{age} = 19.49$, $SD = 1.39$; 69% female; 72.8% White/Caucasian, 10.6% Asian, 9.2% Black/African American, and 7.4% Other).[1] We identified 22 items that were correlated at least |.30| with one of the empathy subscales. From this set of 22 items, we added 8 items identified by an expert sort of a "highly empathic individual" conducted by professional psychologists and advanced graduate students, as reported in Hogan (1969; also see Johnson, 1990). From this set of 30 items, the first and second authors examined each item to ensure that it conceptually tapped into empathy as a construct or at least a close correlate of empathy. We used these considerations to trim the scale to 14 items.

Table 1. Summary of Study Designs and Sample Sizes.

Approximate Birth Year	Block 1969	Guidance 1929	Growth 1929	Oakland 1923	Mills 1936/1938	Radcliffe 1943	Total Observations at Each Wave
Age							*n*
13		73	39	97			209
14	106						106
16		66	40	94			200
18	104						104
21					56		56
23	103						103
30		133					133
36			50				50
40		136		99			235
43					105	103	208
50				97			97
52		142	50				192
53						109	109
60				91			91
62						99	99
73						109	109
Total, *n*	107	171	62	121	112	167	

Note. *N*s are based on number of individuals with at least one assessment of empathy and thus do not sum to the total *N* for each sample.

Table 2. Q-Sort Empathy Items.

Number on Q-Sort	Item
2	I am dependable and responsible person.
5	I am giving and generous toward others.
17	I am a sympathetic and considerate person.
27	I am condescending toward others and act superior to them. (R)
28	Other people tend to like and accept me.
29	People turn to me for advice and reassurance.
35	I am a warm and compassionate person.
36	I tend to undermine, obstruct, or sabotage other people. (R)
37	I sometimes take advantage of other people. (R)
38	I feel a general hostility toward other people. (R)
44	I often try to figure out why people are behaving the way they do.
70	I consider myself an ethical person.
95	I often give advice to other people.
100	I tend to treat everyone the same way.

Note. (R) = reverse scored.

Table 2 presents the final 14 items used in this study that (a) were correlated with the IRI and (b) we thought most closely conceptually represented empathy. Although the minimum average partial (MAP; Garrido, Abad, & Ponsoda, 2011; Velicer, 1976) method suggested two factors, these factors were not interpretable (many items also cross loaded on both factors in an exploratory factor analysis), which led us to conclude the 14 items could most reasonably come to represent a single construct/factor of empathy (see Supplementary Materials).[2] Using the same data, we examined how well our new composite measure of empathy ($\alpha = .86$) correlated with the IRI compared to how well items identified by Hogan (1969) correlated with the IRI. Hogan (1969) first identified 5 items that were most characteristic of an empathic person and then expanded this to 13 items that were "quite characteristic of an empathic person" that included cognitive and emotional items to produce a single score. As seen in Table 3, our CAQ Empathy was more correlated with the IRI (*r*s ranged from .51 to .69) than with Hogan's (1969) items (*r*s ranged from .48 to .72), while the IRI and Hogan items were the least correlated (*r*s ranged from .27 to .50). Thus, we felt that we were reliably measuring the construct of "empathy," which enabled us to examine how this composite changed in the combined data sets previously described.

Results

At each assessment point within each sample, a composite of empathy was constructed ($\alpha_{mean} = .86$; see Supplementary Table S1, for reliabilities at each assessment).[3]

Test–retest (i.e., stability) correlations of empathy are provided in Table 4 ($N = 30$ test–retest correlations). Across all intervals and ages, empathy showed a moderate level of stability over time; however, test–retest correlations depend on the interval and age and reflected autoregressive properties (e.g., $r_{mean} = .479$ for a 10-year interval, $r_{mean} = .371$ for a 20-year interval). As the interval between two assessment points increased, the stability of empathy decreased ($r = -.73$, $p < .001$), replicating work on other psychological characteristics (e.g., Big Five personality traits; Fraley & Roberts, 2005). Further, the stability of empathy was higher among older participants ($r = .42$, $p = .02$), replicating work on rank-order stability of personality (Roberts & DelVecchio, 2000). Given that women tended to score higher in empathy both here and

Table 3. Comparison of Q-Sort Empathy Scales With the Interpersonal Reactivity Index.

Empathy Scales	1	2	3	4	5	*M*	*SD*
(1) Hogan Scale (5 items)						6.32	1.01
(2) Hogan Scale (13 items)	.834					6.56	0.78
(3) Q-Sort Empathy Scale (14 items)	.481	.722				6.82	0.98
(4) IRI empathic concern	.265	.469	.659			3.85	0.68
(5) IRI perspective taking	.324	.434	.509	.457		3.63	0.65
(6) IRI total empathy	.344	.503	.687	.862	.845	3.74	0.57

Note. $N = 783$. IRI = Interpersonal Reactivity Index.
All correlations are significant at $p < .001$.

Table 4. Rank-Order Stability in Empathy.

Gender/ Empathy	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
(1) Gender																
(2) Empathy 13	.128															
(3) Empathy 14	.205*															
(4) Empathy 16	−.002	.540**														
(5) Empathy 18	.243*		.572**													
(6) Empathy 21																
(7) Empathy 23	.284**		.554**		.622**											
(8) Empathy 30	−.003	.237*		.330**												
(9) Empathy 36	.329*	.293		.358*												
(10) Empathy 40	.194**	.260**		.183*				.260**								
(11) Empathy 43						.293*										
(12) Empathy 50	.131	.271*		.076						.461**						
(13) Empathy 52	.228**	.212*		.325**				.327**	.345*	.309**						
(14) Empathy 53											.684**					
(15) Empathy 60	.249*	.327**		.084						.384**		.466**				
(16) Empathy 62											.405**			.591**		
(17) Empathy 72											.298*			.481**		.507**

Note. Empathy 13 = empathy at age 13.
*$p < .05$. **$p < .01$.

in prior work (Davis, 1983; O'Brien et al., 2013), participant gender was included as a covariate in all analyses.

Main Analysis

The six samples were combined for the purposes of multilevel analyses in a variant of an accelerated longitudinal design, in which data sets with overlapping ages are combined to estimate growth curves across the entirety of the data (Helson et al., 2002; Hirschberger, Srivastava, Marsh, Cowan, & Cowan, 2009; Miyazaki & Raudenbush, 2000; Raudenbush & Chan, 1992; Terracciano, McCrae, Brant, & Costa, 2005). Following recommendations that it is more appropriate to examine research questions within one larger sample rather than smaller, separate samples (Schimmack, 2012), we combined the samples because of the similarity and overlap between them, to maximize power and present the data in the most parsimonious way. We present the results from this combined sample. Sample source was added as a covariate in the analyses reported below. The Radcliffe sample (the eldest sample) was added as the reference group against which the other samples were compared.

Prior to computing the interaction terms, age was grand-mean centered, and gender was contrast coded ($-1 = men$, $1 = women$) and treated as time invariant. Empathy was predicted from age, age^2, gender, and the interaction between these variables. The inclusion of quadratic effects of age was based on previous research that had found quadratic effects of age on empathy (O'Brien et al., 2013). For the main analyses, we did not model more complex terms as there were no theoretical reasons to expect a developmental process to follow a cubic (or more complicated) age trend across the life span.[4]

To examine whether and how empathy changed across the life span, we used growth curve modeling in the SPSS 22 Mixed command (Peugh & Enders, 2005), which enabled us to model intraindividual changes and moderators of these changes (Roberts & Chapman, 2000; Willett, 1988). Degrees of freedom were computed using the Satterthwaite's (1946) method. The variance of the intercepts and slopes (of age and

Table 5. Multilevel Model Predicting Changes in Empathy Across the Life Span.

Predictors	*b*	*SE*	*df*	β	*t*	*p*	LB	UB
Age	.02	.002	1,767.03	.38	12.16	<.001	.02	.02
Gender	.27	.04	1,281.99	.13	6.54	<.001	.19	.35
Age × Gender	−.002	.002	1,677.33	−.03	−0.96	.337	−.01	.002
Age^2	<.001	<.001	1,970.48	.08	2.64	.008	<.001	.001
Age^2 × Gender	<.001	<.001	1,992.21	−.12	−4.27	<.001	−.001	<.001
Block	.42	.11	1,072.64		3.70	<.001	.20	.65
Berkeley Guidance	−.28	.09	1,033.99		−3.01	.003	−.46	−.10
Berkeley Growth	−.11	.11	954.35		−0.98	.329	−.33	.11
Oakland Growth	−.28	.09	965.06		−2.96	.003	−.46	−.09
Mills	−.19	.11	1,299.99		−1.81	.07	−.40	.02

Note. Radcliffe College sample is the reference group for the dummy-coded sample source variable. Gender: −1 = *men*, 1 = *women*. Age was grand-mean centered prior to analysis.

age^2) were significant (*p*s < .001), suggesting significant individual differences in the effects of age on empathy.

Results from the growth curve analysis for empathy are presented in Table 5. The effects of age, gender, and age^2 were significant predictors of empathy. Empathy increased on average across the life span (Figure 1). The largest increases in empathy occurred after age 40. Women were higher in empathy on average. An Age^2 × Gender interaction emerged as significant. Decomposing this interaction revealed that the quadratic effect was more pronounced for men (β = .24, $p < .001$) than women (β = −.05, $p = .02$). However, because the eldest participants in the sample were all women (i.e., the Radcliffe and Mills samples), drawing comparisons between men and women is difficult. Thus, we will not devote considerable space to interpreting these differences given the unbalanced gender design of the study.

Relative to the Radcliffe College sample (i.e., the eldest participants), participants in the other samples were generally lower in empathy. After controlling for the effects of age and gender, participants from the Block and Block sample were higher in empathy. However, this may obscure our direct test of whether empathy has been increasing or decreasing in recent cohorts. Although we entered in cohort/sample as a blocking variable, we also recoded the variable to represent the sample's birth year[5] and modeled the effect as a supplementary analysis. Results suggested that empathy is slightly increasing in recent birth cohorts, $b = .02$, $SE = .002$, $t(803.81) = 8.54$, $p < .001$. Given the possibility of a two-factor empathy measure (see above), we tested whether the main results differed systematically across these two factors. We found similar results across both factors of empathy (i.e., both factors of empathy increased over time; see Supplementary Materials for details).

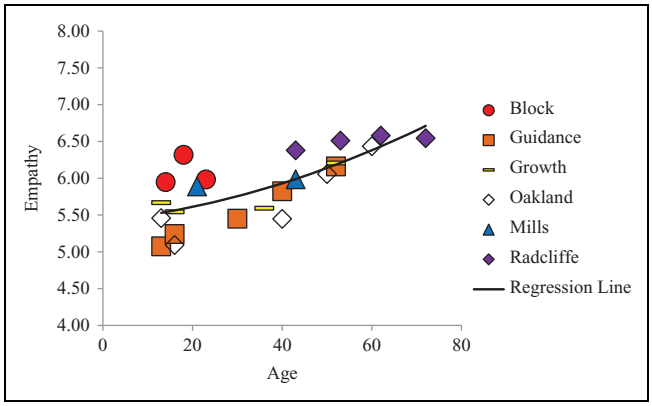

Figure 1. Longitudinal changes in empathy across the life span.

Discussion

While previous research showed mixed results with respect to the direction and magnitude of empathy changes, we found that empathy increases from adolescence to older adulthood using observer ratings of empathy. The current study's use of observer ratings provides support for this pattern of change that has been found occasionally in previous research. Throughout adolescence and early adulthood, people learn to think consider perspectives and emotions that might differ from their own (Davis & Franzoi, 1991; Sze et al., 2012). The pursuit of contexts that challenge one's perspective is an important characteristic of this life stage, which possibly drives the empathy development (Arnett, 2000). Our findings are also consistent with older adults' higher self-reports of emotional empathy and prosocial behavior (Sze et al., 2012). Further, as people age, they start to perceive their remaining time as limited, leading to a heightened focus on achieving emotional goals that improve well-being, and higher empathy is closely related to these goals (Carstensen, 2006; Carstensen, Fung, & Charles, 2003; Carstensen et al., 1999). Finally, consistent with Grühn, Rebucal, Diehl, Lumley, and Labouvie-Vief (2008), we found some support for generational increases in empathy, suggesting that both age- and cohort-related changes in empathy are occurring. However, we do not have the exact age overlap in each sample, and some samples do not include enough ages to make this cohort/sample effect a perfect comparison of different cohorts.

Identifying the mechanisms of change is particularly important for revealing the increase in empathy in middle age and older adulthood. What happens during this time that

precipitates change? We have discussed some of the psychological changes that occur after young adulthood that might affect people's empathy (e.g., changes in socioemotional goals). However, attention has also been paid to examining how significant life events and more structured changes alter people's emotions and development over time (Jokela, Kivimaki, Elovainio, & Keltikangas-Jarvinen, 2009; Löckenhoff, Terracciano, & Costa, 2009; Löckenhoff, Terracciano, Patriciu, Eaton, & Costa, 2009; Mangelsdorf, Eid, & Luhmann, 2018; Specht et al., 2014; Specht, Egloff, & Schmukle, 2011). For example, having children, witnessing the declining health of family members, or reflecting on one's place in the world after retirement may change people's perspectives on life and trigger more empathic processes. We consider this an important direction for future research.

Limitations

There are limitations to the current study that are worth acknowledging. Based on the results from Table 3, the CAQ Empathy measure may more accurately capture the emotional empathy at the expense of cognitive empathy. Although some items clearly involve cognitive empathy (e.g., "I often try to figure out why people are behaving the way they do"), there was not enough items to reliably assess cognitive empathy via the CAQ. The MAP test suggested a two-factor solution for the empathy measure, which may roughly approximate this emotional and cognitive distinction; indeed, the aforementioned item falls under a factor along with some items related to advice giving, which might be a marker of cognitive empathy that reflects perspective taking ability. Supplementary analyses suggested both factors of empathy increased across the life span. However, these results do not necessarily imply that both emotional and cognitive empathy increase. Rather, given the items comprising both factors, our efforts suggest that the CAQ-Empathy measure is closer to a unidimensional measure, focusing primarily on *emotional* empathy. Considering these psychometric properties, our results may not perfectly correspond to studies that more formally distinguish between emotional and cognitive empathy. Yet we believe that the CAQ-Empathy Scale can still be a tool that addresses some practical challenges in data collection (e.g., requiring a measure of a certain construct after completing data collection), albeit it might be more related to emotional empathy. Future researchers using this measure therefore should interpret their findings with caution, acknowledging that their results likely reflect emotional empathy and may not perfectly generalize to studies using other empathy measures (e.g., perspective taking in IRI). Additional longitudinal data using instruments that measure different facets of empathy are needed to provide a more comprehensive look at how empathy changes over time.

Further, although we combined samples from different studies and analyzed them as if they were one contiguous sample, they are indeed different samples of participants followed over different intervals of time. Thus, participants were not followed over the same ages, and we do not have the full data for any one individual from age 13 to 72. In this way, our study resembles other work in which different groups of individuals are followed over time and developmental interpretations are made on the basis of these data (Grühn et al., 2008; Terracciano et al., 2005).

There are several measures that we would have ideally included to more appropriately measure not only empathy (i.e., separate components) but also the mechanisms of changes in empathy. We could not see whether there are different developmental courses for each component of empathy and could only speculate about the underlying causes of change. We hope that future research will more formally examine why empathy changes across the life span (i.e., mediators) and the conditions and contexts under which these changes are magnified/diminished (i.e., moderators). There are also methodological questions that are left unanswered by this study. For example, CAQs were completed by a different composition of raters across and within studies. Unfortunately, not enough information is known about the particular raters (e.g., if the rater was a student or clinician) to formally model the ways in which this might have affected the results. Also, in many cases, the types of informants (e.g., clinicians, research personnel) were collinear with sample source (i.e., some studies only employed research personnel). Nevertheless, the test–retest correlations were comparable to other studies of individual differences of self-reports and single observer reports. This stability suggests that raters have reliably completed CAQs across waves, even though the composition of raters and underlying participant information available likely changed over time. Future research can examine whether estimates of change in empathy are different when assessed via self-reports and from varying informants (Watson & Humrichouse, 2006).

Conclusion

The current study combined data from six samples to examine age-related differences in empathy from age 13 to 72. Contrary to research finding declines or high stability in empathy, empathy continued to increase across the life span, controlling for birth cohort. Empathy has many intra- and interpersonal benefits for individuals, and identifying the conditions under which empathy can be cultivated across the life span is an important direction for future research.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Supplemental Material

The supplemental material is available in the online version of the article.

Notes

1. The smallest possible effect we could detect with this sample was Cohen's $f^2 = .01$ (at 80% power, $\alpha = .05$).
2. We consider the use of an existing measure (the California Adult Q-Sort) and the single dimension of empathy to be an unfortunate limitation to using secondary data and discuss the limitations of the measure at length in the Discussion section.
3. There are different ways of assessing invariance with a study design like ours; these methods vary in complexity and exactly how they test for different types of invariance. In discussing it among the author team, we considered testing invariance in many ways, ranging from traditional methods (e.g., confirmatory factor analyses [CFAs]), variants of more atypical methods (e.g., multitrait multimethod approaches), or specifically focusing on the "linkages" or places where the time points across data sets overlap to ensure comparability on the empathy measure. The uneven design (e.g., participants from different samples being assessed at different ages/intervals) precluded our ability to employ succinct approaches to conduct invariance tests across the entirety of the data (e.g., multilevel CFAs in which assessments can be nested within individuals or samples; Geldhof, Preacher, & Zyphur, 2014). Because of the large follow-up periods within each individual data set, we felt it was appropriate to employ two approaches to testing measurement invariance. First, we examined configural, metric, and scalar invariance within each sample over time (to assess longitudinal invariance). Testing each form of invariance in a sequential fashion revealed that empathy demonstrated configural, metric, and scalar invariance within each of the samples (root mean square error of approximations [ΔRMSEAs] < .02; see Nye & Drasgow, 2011, for a discussion of effect sizes for invariance tests in the context of CFA). Second, we examined configural, metric, and scalar invariance between adjacent samples at the closest assessment points (to ensure that empathy's measurement and systematic variability was comparable across samples). For example, we compared the age 43 assessment point of the Mills sample with the age 43 assessment point of the Radcliffe sample. Likewise, the age 16 assessment points of the Berkeley Guidance, Berkeley Growth, and Oakland Growth studies were compared. Results from these multigroup CFAs also revealed strong configural, metric, and scalar invariance (ΔRMSEAs < .02). In the few cases in which scalar invariance was not achieved (i.e., between the age 23 Block assessment and the age 21 Mills sample), departures from invariance were not extreme (e.g., ΔRMSEAs = .04), and the models still demonstrated acceptable fit. We can thus reasonably conclude that our measure of empathy was largely invariant over time and across samples. However, there was no one approach to testing measurement invariance in a design such as ours (C. D. Nye, personal communication, October 1, 2018). Additional methods for testing invariance in contexts such as the design of our study should be the subject of future inquiry.
4. We also elected to try to fit a latent basis model to the data using MPlus 8.1. A latent basis model estimates individual slopes based on the observations rather than an a priori model of change (e.g., quadratic). Interestingly, the latent basis model performed slightly worse compared to the quadratic model (ΔRMSEA = .03, Δ comparative fit index = .105). Because of the worse fit of the latent basis model and precedent set from previous research (O'Brien, Konrath, Grühn, & Hagen, 2013), the quadratic model was chosen for the final analyses.
5. The birth year for the Mills sample was coded as 1937 since they could have been born either in 1936 or in 1938. Birth year was grand-mean centered prior to analysis.

References

Allemand, M., Steiger, A. E., & Fend, H. A. (2015). Empathy development in adolescence predicts social competencies in adulthood. *Journal of Personality*, *83*, 229–241. doi:10.1111/jopy.12098

Arnett, J. J. (1998). Learning to stand alone: The contemporary American transition to adulthood in cultural and historical context. *Human Development*, *41*, 295–315.

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, *55*, 469–480. doi:10.1037/0003-066x.55.5.469

Bailey, P. E., & Henry, J. D. (2008). Growing less empathic with age: Disinhibition of the self-perspective. *Journals of Gerontology Series B Psychological Sciences and Social Sciences*, *63*, 219–226.

Bailey, P. E., Henry, J. D., & Von Hippel, W. (2008). Empathy and social functioning in late adulthood. *Aging & Mental Health*, *12*, 499–503. doi:10.1080/13607860802224243

Beadle, J. N., Sheehan, A. H., Dahlben, B., & Gutchess, A. H. (2015). Aging, empathy, and prosociality. *Journals of Gerontology Series B Psychological Sciences and Social Sciences*, *70*, 215–224. doi:10.1093/geronb/gbt091

Block, J. (1961). *The Q-sort method in personality assessment and psychiatric research*. Springfield, IL: Charles C Thomas.

Block, J. (1971). *Lives through time*. Berkeley, CA: Bancroft Books.

Block, J. (2008). *The Q-sort in character appraisal: Encoding subjective impressions of persons quantitatively*. Washington, DC: American Psychological Association.

Block, J., & Block, J. H. (2006). Venturing a 30-year longitudinal study. *American Psychologist*, *61*, 315–327. doi:10.1037/0003-066x.61.4.315

Cabello, R., Sorrel, M. A., Fernandez-Pinto, I., Extremera, N., & Fernandez-Berrocal, P. (2016). Age and gender differences in ability emotional intelligence in adults: A cross-sectional study. *Developmental Psychology*, *52*, 1486–1492. doi:10.1037/dev0000191

Carstensen, L. L. (2006). The influence of a sense of time on human development. *Science*, *312*, 1913–1915. doi:10.1126/science.1127488

Carstensen, L. L., Fung, H. H., & Charles, S. T. (2003). Socioemotional selectivity theory and the regulation of emotion in the second half of life. *Motivation and Emotion*, *27*, 103–123. doi:10.1023/a:1024569803230

Carstensen, L. L., Isaacowitz, D. M., & Charles, S. T. (1999). Taking time seriously: A theory of socioemotional selectivity. *American Psychologist*, *54*, 165–181. doi:10.1037/0003-066x.54.3.165

Chen, Y.-C., Chen, C.-C., Decety, J., & Cheng, Y. (2013). Aging is associated with changes in the neural circuits underlying empathy. *Neurobiology of Aging*, *30*, 1–10.

Chopik, W. J. (2017). Associations among relational values, support, health, and well-being across the adult lifespan. *Personal Relationships*, *24*, 408–422. doi:10.1111/pere.12187

Chopik, W., O'Brien, E. H., & Konrath, S. H. (2017). Differences in empathic concern and perspective taking across 63 countries. *Journal of Cross-Cultural Psychology*, *48*, 23–38.

Chopik, W. J., & Edelstein, R. S. (2015). Preliminary validation of a romantic attachment orientation measure from the California Adult Q-Sort. *Attachment and Human Development*, *17*, 570–585. doi: 10.1080/14616734.2015.1091852

Chopik, W. J., Joshi, D. H., & Konrath, S. (2014). Historical changes in self-interest: Presidential addresses from 1790 to 2012. *Personality and Individual Differences*, *66*, 128–133.

Cramer, D., & Jowett, S. (2010). Perceived empathy, accurate empathy and relationship satisfaction in heterosexual couples. *Journal of Social and Personal Relationships*, *27*, 327–349. doi:10.1177/0265407509348384

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*, 113–126. doi:10.1037/0022-3514.44.1.113

Davis, M. H., & Franzoi, S. L. (1991). Stability and change in adolescent self-consciousness and empathy. *Journal of Research in Personality*, *25*, 70–87. doi:10.1016/0092-6566(91)90006-C

Davis, M. H., & Oathout, H. A. (1987). Maintenance of satisfaction in romantic relationships: Empathy and relational competence. *Journal of Personality and Social Psychology*, *53*, 397–410. doi:10.1037/0022-3514.53.2.397

Eichorn, D. H., Clausen, J. A., Haan, N., Honzik, M. P., & Mussen, P. H. (1981). *Present and past in middle life*. New York, NY: Academic Press.

Eisenberg, N., Guthrie, I. K., Murphy, B. C., Shepard, S. A., Cumberland, A., & Carlo, G. (1999). Consistency and development of prosocial dispositions: A longitudinal study. *Child Development*, *70*, 1360–1372. doi:10.1111/1467-8624.00100

Eysenck, S. B., Pearson, P. R., Easting, G., & Allsopp, J. F. (1985). Age norms for impulsiveness, venturesomeness and empathy in adults. *Personality and Individual Differences*, *6*, 613–619. doi: 10.1016/0191-8869(85)90011-X

Fraley, R. C., & Roberts, B. W. (2005). Patterns of continuity: A dynamic model for conceptualizing the stability of individual differences in psychological constructs across the life course. *Psychological Review*, *112*, 60–74. doi:10.1037/0033-295x.112.1.60

Garrido, L. E., Abad, F. J., & Ponsoda, V. (2011). Performance of Velicer's minimum average partial factor retention method with categorical variables. *Educational and Psychological Measurement*, *71*, 551–570.

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods*, *19*, 72.

Gilet, A.-L., Mella, N., Studer, J., Grühn, D., & Labouvie-Vief, G. (2013). Assessing dispositional empathy in adults: A French validation of the Interpersonal Reactivity Index (IRI). *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, *45*, 42.

Greenberg, E. H., & Weber, K. (2008). *Generation we: How millennial youth are taking over America and changing the world*. Emeryville, CA: Pachatusan.

Grühn, D., Rebucal, K., Diehl, M., Lumley, M., & Labouvie-Vief, G. (2008). Empathy across the adult lifespan: Longitudinal and experience-sampling findings. *Emotion*, *8*, 753–765. doi:10.1037/a0014123

Haan, N., Millsap, R., & Hartka, E. (1986). As time goes by: Change and stability in personality over fifty years. *Psychology and Aging*, *1*, 220–232.

Hannikainen, I. R., Machery, E., & Cushman, F. A. (2018). Is utilitarian sacrifice becoming more morally permissible? *Cognition*, *170*, 95–101. doi:10.1016/j.cognition.2017.09.013

Helson, R. (1993). The Mills classes of 1958 and 1960: College in the fifties, young adulthood in the sixties. In K. D. Hulbert & D. T. Schuster (Eds.), *Women's lives through time* (pp. 190–210). San Francisco, CA: Jossey-Bass.

Helson, R., Jones, C., & Kwan, V. S. Y. (2002). Personality change over 40 years of adulthood: Hierarchical linear modeling analyses of two longitudinal samples. *Journal of Personality and Social Psychology*, *83*, 752–766. doi:10.1037/0022-3514.83.3.752

Hirschberger, G., Srivastava, S., Marsh, P., Cowan, C. P., & Cowan, P. A. (2009). Attachment, marital satisfaction, and divorce during the first fifteen years of parenthood. *Personal Relationships*, *16*, 401–420.

Hoffman, M. L. (2008). Empathy and prosocial behavior. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (Vol. *3*, pp. 440–455). New York, NY: Guilford Press.

Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology*, *33*, 307–316.

Howe, N., & Strauss, B. (2000). *Millennials rising: The next great generation*. New York, NY: Vintage.

Johnson, J. A. (1990). Empathy as a personality disposition. In R. C. Mackay, J. R. Hughes, & E. J. Carver (Eds.), *Empathy in the helping relationship* (pp. 49–64). New York, NY: Springer.

Jokela, M., Kivimaki, M., Elovainio, M., & Keltikangas-Jarvinen, L. (2009). Personality and having children: A two-way relationship. *Journal of Personality and Social Psychology*, *96*, 218–230. doi: 10.1037/a0014058

Konrath, S. H., O'Brien, E., & Hsing, C. (2011). Changes in dispositional empathy in American college students over time: A meta-analysis. *Personality and Social Psychology Review*, *15*, 180–198. doi:10.1177/1088868310377395

Labouvie-Vief, G., Grühn, D., & Studer, J. (2010). Dynamic integration of emotion and cognition: Equilibrium regulation in development and aging. In R. M. Lerner, M. E. Lamb, & A. M. Freund (Eds.), *The handbook of life-span development* (pp. 79–115). Hoboken, NJ: Wiley.

Löckenhoff, C. E., Terracciano, A., & Costa, P. T. (2009). Five-factor model personality traits and the retirement transition: Longitudinal and cross-sectional associations. *Psychology and Aging*, *24*, 722–728. doi:10.1037/a0015121

Löckenhoff, C. E., Terracciano, A., Patriciu, N. S., Eaton, W. W., & Costa, P. T. (2009). Self-reported extremely adverse life events and longitudinal changes in five-factor model personality traits in an urban sample. *Journal of Traumatic Stress*, *22*, 53–59. doi:10.1002/jts.20385

Mangelsdorf, J., Eid, M., & Luhmann, M. (2018). Does growth require suffering? A systematic review and meta-analysis on genuine

posttraumatic and postecstatic growth. *Psychological Bulletin, 145*, 302–338. doi:10.1037/bul0000173

Miyazaki, Y., & Raudenbush, S. W. (2000). Tests for linkage of multiple cohorts in an accelerated longitudinal design. *Psychological Methods*, *5*, 44–63.

Newton, N. J., & Stewart, A. J. (2013). The road not taken: Women's life paths and gender-linked personality traits. *Journal of Research in Personality*, *47*, 306–316. doi:10.1016/j.jrp.2013.02.003

Nye, C. D., & Drasgow, F. (2011). Effect size indices for analyses of measurement equivalence: Understanding the practical importance of differences between groups. *Journal of Applied Psychology*, *96*, 966–980. doi:10.1037/a0022955

O'Brien, E., Konrath, S. H., Grühn, D., & Hagen, A. L. (2013). Empathic concern and perspective taking: Linear and quadratic effects of age across the adult life span. *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*, *68B*, 168–175. doi:10.1093/geronb/gbs055

Peugh, J. L., & Enders, C. K. (2005). Using the SPSS mixed procedure to fit cross-sectional and longitudinal multilevel models. *Educational and Psychological Measurement*, *65*, 717–741. doi:10.1177/0013164405278558

Raudenbush, S. W., & Chan, W.-S. (1992). Growth curve analysis in accelerated longitudinal designs. *Journal of Research in Crime and Delinquency*, *29*, 387–411.

Rifkin, J. (2010). *The empathic civilization: The race to global consciousness in a world in crisis*. New York, NY: Tarcher.

Roberts, B. W., & Chapman, C. N. (2000). Change in dispositional well-being and its relation to role quality: A 30-year longitudinal study. *Journal of Research in Personality*, *34*, 26–41. doi:10.1006/jrpe.1999.2259

Roberts, B. W., & Mroczek, D. (2008). Personality trait change in adulthood. *Current Directions in Psychological Science*, *17*, 31–35.

Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. *Neuroscience & Biobehavioral Reviews*, *32*, 863–881.

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin*, *2*, 110–114. doi:10.2307/3002019

Schimmack, U. (2012). The ironic effect of significant results on the credibility of multiple-study articles. *Psychological Methods*, *17*, 551–566. doi:10.1037/a0029487

Specht, J., Bleidorn, W., Denissen, J. J. A., Hennecke, M., Hutteman, R., Kandler, C., . . . Zimmermann, J. (2014). What drives adult personality development? A comparison of theoretical perspectives and empirical evidence. *European Journal of Personality*, *28*, 216–230. doi:10.1002/per.1966

Specht, J., Egloff, B., & Schmukle, S. C. (2011). Stability and change of personality across the life course: The impact of age and major life events on mean-level and rank-order stability of the Big Five. *Journal of Personality and Social Psychology*, *101*, 862–882. doi:10.1037/a0024950

Stewart, A. J., & Vandewater, E. A. (1993). The Radcliffe class of 1964: Career and family social clock projects in a transitional cohort. In K. D. Hulbert & D. T. Schuster (Eds.), *Women's lives through time: Educated American women of the twentieth century* (pp. 235–258). San Francisco, CA: Jossey-Bass.

Sze, J. A., Gyurak, A., Goodkind, M. S., & Levenson, R. W. (2012). Greater emotional empathy and prosocial behavior in late life. *Emotion*, *12*, 1129–1140. doi:10.1037/a0025011

Terracciano, A., McCrae, R. R., Brant, L. J., & Costa, P. T. (2005). Hierarchical linear modeling analyses of NEO-PI-R scales in the Baltimore Longitudinal Study of Aging. *Psychology and Aging*, *20*, 493–506. doi:10.1037/0882-7974.20.3.493

Twenge, J. M. (2006). *Generation me: Why today's young Americans are more confident, assertive, entitled—And more miserable than ever before*. New York, NY: Free Press.

Van der Graaff, J., Branje, S., De Wied, M., Hawk, S., Van Lier, P., & Meeus, W. (2016). Perspective taking and empathic concern in adolescence: Gender differences in developmental changes. *Developmental Psychology*, *50*, 881–888. doi:10.1037/a0034325

Vazire, S., & Carlson, E. N. (2011). Others sometimes know us better than we know ourselves. *Current Directions in Psychological Science*, *20*, 104–108. doi:10.1177/0963721411402478

Vazire, S., & Mehl, M. R. (2008). Knowing me, knowing you: The accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. *Journal of Personality and Social Psychology*, *95*, 1202–1216.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*, 321–327.

Watson, D., & Humrichouse, J. (2006). Personality development in emerging adulthood: Integrating evidence from self-ratings and spouse ratings. *Journal of Personality and Social Psychology*, *91*, 959–974.

Willett, J. B. (1988). Questions and answers in the measurement of change. *Review of Research in Education*, *15*, 345–422.

Zhang, X., Fung, H. H., Stanley, J. T., Isaacowitz, D. M., & Ho, M. Y. (2013). Perspective taking in older age revisited: A motivational perspective. *Developmental Psychology*, *49*, 1848.

Author Biographies

Jeewon Oh, BA, is a graduate student at Michigan State University. She studies family and close relationship process across the life span and how these processes are associated with intra- and interpersonal adjustment.

William J. Chopik, PhD, is an assistant professor at Michigan State University. He studies the continuity and change of relationship processes across the life span and individual differences in responses to intimacy.

Sara Konrath, PhD, is an associate professor at Indiana University–Purdue University Indianapolis. She studies the costs and benefits of empathy and related traits (e.g., emotional intelligence, narcissism) and behaviors (e.g., helping, caregiving). She also examines causes, consequences, and changes in such topics from a variety of perspectives, including psychophysiological and neurological.

Kevin J. Grimm, PhD, is a professor at Arizona State University. His research focuses on longitudinal methods for the study of change at the individual and group level.

Handling Editor: Maike Luhmann