

3 Method

3.1 Participants

3.1.1 *Sample Size and Composition*

The original sample comprised 316 participants with an age range of 19 to 81 years. A survey company recruited the majority of participants by means of a random dialing procedure in the city of Berlin (Germany). Because the company did not always reach the desired number of persons with particular background characteristics, additional recruitment strategies were used. Sixteen individuals who had participated in unrelated studies (e.g., on stereotypes, cognitive training, and word perception) and 18 individuals who were acquainted with study participants from earlier studies were contacted by phone and asked to participate in the study. Finally, four older individuals were recruited during visits to senior centers in the city of Berlin. Upon recruitment, participants were informed that the present study investigated the nature of personal longings and what persons think and feel when they experience longing. Participants received 50 Euro for their participation in three test sessions.

Ten participants had to be excluded because they failed to provide usable data (7 participants did not report any longings and 3 participants failed to fill out more than half of the questionnaires). In addition, 7 participants were randomly excluded from over-recruited cells to reduce differences in cell composition (resulting in a range of 10 to 15 participants per cell)³. This resulted in an effective baseline sample of 299 participants. Table 3 contains the sample composition. The sample was stratified by age (six age blocks of ten years), sex⁴ (51 % men and 49 % women), and education (44 % low and 56 % high)⁵. For several analyses, the sample was

³ Over-recruitment occurred primarily in high education cells. It cannot be ruled out that persons with different degrees of education differ in the structure of longing or other central variables in this study. Thus, overrepresentation of persons with high education may bias results. Limiting the size of cells to a range of 10 to 15 participants attempted to control for this possibility.

⁴ Throughout the dissertation, the term “sex” rather than “gender” will be used to refer to differences between men and women. This does not imply that biological factors are considered to be the exclusive or even primary causes for potential differences between men and women. Mostly, psychological and social factors may be equally or even more important. Nevertheless, the term “sex” appeared to be more appropriate given the current operationalization. No information was obtained about gender-related aspects of the self-concept (e.g., masculinity and femininity).

⁵ High education meant that participants had completed the highest track of German school education (12 to 13 years excluding kindergarten; in German: Abitur).

divided into three age groups: young adults (19 to 39 years; $n = 98$), middle-aged adults (40 to 59 years; $n = 102$), and old adults (60 to 81 years; $n = 99$).

Table 3
Composition of the Baseline Sample

Age groups		Men		Women		Σ
		Low education	High education	Low education	High education	
Young adults	19 – 29 years	10	15	11	12	48
	30 – 39 years	10	13	13	14	50
Middle-aged adults	40 – 49 years	11	14	11	15	51
	50 – 59 years	13	14	10	14	51
Old adults	60 – 69 years	12	15	11	15	53
	70 – 81 years	10	15	11	10	46
	Σ	66	86	67	80	$N = 299$

Note. Participants with high education had completed the highest track of German school education (12 to 13 years excluding kindergarten; in German: Abitur).

3.1.2 *Socio-Demographic Characteristics*

Table 4 gives an overview of the socio-demographic characteristics of the total sample and for the young, middle-aged, and old subsamples. Although the sample was stratified for education, it can be seen that in the group of highly educated individuals (high school degree or higher), participants with a college or university degree were overrepresented. Regarding family status, more than half of the total sample reported to live in a partnership (long-term partnership or married) and about two-thirds of the sample reported to have children. Regarding work status, about one third of the sample was part-time or full-time employed, one third was retired, and the remaining third was still in the process of education, unemployed, or homemaker.

Age group differences in socio-demographic characteristics were largely comparable to the general population (Federal Statistical Office, 2002). There was, however, one notable exception. Whereas in the general population, the number of individuals with a college or university degree is lower in older age cohorts, the opposite was true for the present sample. Consequently, education (with the four levels shown in Table 4) was included as a control variable in central analyses.

Table 4
Socio-Demographic Characteristics of the Total Sample and for Subsamples of Young, Middle-Aged, and Old Adults

Variables	Total sample (19 to 81 years)		Young adults (19 to 39 years)		Middle-aged adults (40 to 59 years)		Old adults (60 to 81 years)	
	N = 299		n = 98		n = 102		n = 99	
<i>Age</i>								
Mean (SD)	49.9	(17.0)	31.2	(6.0)	49.9	(6.3)	69.4	(5.3)
<i>Sex</i>								
Female	147	49.2 %	50	51.0 %	50	49.0 %	47	47.5 %
Male	152	50.8 %	48	49.0 %	52	51.0 %	52	52.5 %
<i>Education</i>								
Primary education (8-9 years) ^a	45	15.0 %	7	7.1 %	15	14.7 %	23	23.2 %
Lower secondary education (10 years) ^b	88	29.4 %	37	37.8 %	30	29.4 %	21	21.2 %
High school (12-13 years) ^c	50	16.7 %	35	35.7 %	9	8.8 %	6	6.1 %
College/University ^d	116	38.8 %	19	19.4 %	48	47.1 %	49	49.5 %
<i>Marital Status</i>								
Single	53	17.7 %	34	34.7 %	12	11.8 %	7	7.1 %
Long-term partnership	56	18.7 %	43	43.9 %	10	9.8 %	3	3.0 %
Married	115	38.5 %	16	16.3 %	46	45.1 %	53	53.5 %
Divorced	49	16.4 %	4	4.1 %	29	28.4 %	16	16.2 %
Widowed	26	8.7 %	1	1.0 %	5	4.9 %	20	20.2 %
<i>Number of Children</i>								
None	101	33.8 %	69	70.4 %	20	19.6 %	12	12.1 %
One child	57	19.1 %	10	10.2 %	20	19.6 %	27	27.3 %
Two children	95	31.8 %	13	13.3 %	45	44.1 %	37	37.4 %
More than two children	43	14.3 %	4	4.1 %	16	15.7 %	23	23.2 %
<i>Current Occupation</i>								
Full-time employed	64	21.4 %	21	21.4 %	36	35.3 %	7	7.1 %
Part-time employed	36	12.0 %	19	19.4 %	16	15.7 %	1	1.0 %
Unemployed	43	14.4 %	18	18.4 %	25	24.5 %	0	0.0 %
Retired	101	33.8 %	0	0.0 %	12	11.8 %	89	89.9 %
Student	24	8.0 %	23	23.5 %	1	1.0 %	0	0.0 %
High school student or job trainee	11	3.7 %	11	11.2 %	0	0.0 %	0	0.0 %
Homemaker	13	4.3 %	5	5.1 %	6	5.9 %	2	2.0 %

Note. ^a German: Volks-/ Gemeinde-/ Hauptschule. ^b German: Mittlere Reife / Realschule. ^c German: (Fach-) Abitur. ^d German: Fach-/ Hochschulstudium.

3.1.3 Self-Rated Health and Intellectual Performance

Self-rated physical and mental health. Three items were used to assess the self-rated physical and mental health status of the sample: (1) “How would you rate your physical health at

present?'' (2) ''How would you rate your mental health at present?'' and (3) ''How would you rate your mental fitness at present?'' Responses were given on a six-point scale ranging from 0 (*very poor*) to 5 (*excellent*). As seen in Table 5, on average, participants rated their physical and mental health as moderate to good. Age was associated with poorer physical health, although the effect was only small. No significant age associations emerged with respect to self-rated mental health and fitness. Overall, in terms of self-rated health, the sample was comparable to previous studies (e.g., J. Heckhausen et al., 1989; Salthouse & Ferrer-Caja, 2003).

Table 5
Self-Rated Health and Intellectual Performance: Descriptive Information and Correlation With Age

Constructs	<i>M</i>	<i>SD</i>	<i>r</i> _{Age} (<i>p</i>)
<i>Self-rated health</i>			
Self-rated physical health	3.43	1.12	-.17 (.001)*
Self-rated mental health	3.62	1.13	.10 (.09)
Self-rated mental fitness	3.81	.88	-.04 (.52)
<i>Intellectual performance</i>			
Mental Speed	48.79	11.33	-.58 (.001)*
Verbal knowledge	26.36	4.27	.42 (.001)*

Note. * $p < .01$

Intellectual performance. Participants completed screening measures for fluid and crystallized intelligence (P. B. Baltes, 1987; Cattell, 1971; Horn, 1982). These were the Digit-Symbol-Coding-Test (Wechsler, 1955) for mental speed and the Spot-A-Word-Test (adapted from Lehrl, 1977) for verbal knowledge (see Lindenberger, Mayr, & Kliegl, 1993 for a description of both tests). Results are also shown in Table 5.

Consistent with previous research (e.g., P. B. Baltes & Lindenberger, 1997; Li et al., 2004; Salthouse & Ferrer-Caja, 2003), age was negatively correlated with processing speed. The strength of association was comparable to most previous studies. Verbal knowledge, in contrast, was positively correlated with age. Previous studies with a comparable age range also found a positive, but less steep, age trajectory for verbal knowledge (e.g., Li et al., 2004; Verhaeghen, 2003). Taken together with the high percentage of middle-aged and older participants with a college or university degree, this finding suggests that older participants were more positively selected than younger participants. Including education level as control variable in central analyses attempted to control for this bias.

3.2 Procedure

The majority of participants attended three group sessions of 3 to 20 people ($M = 12$ participants) that were run by trained research assistants. Groups were mixed with regard to socio-demographic characteristics (i.e., age, sex, and education). Fourteen additional appointments with one or two persons were made for participants who were unable to attend regular sessions or missed their initial appointment. For these participants, a special effort was made to ensure that the time interval between the first and third sessions was approximately five weeks. The first and third test sessions lasted on average 100 minutes, whereas the second session lasted on average 65 minutes.

Longing was assessed in the first session by means of a “guided mental journey” through different life phases, followed by the completion of a questionnaire (for details, see below). In the second session (on average 9 days after the first session; $SD = 6.7$ days), participants filled out a number of questionnaires designed to assess subjective well-being, personality, generalized control beliefs, self-regulation, and intelligence. Only data on intelligence, generalized control beliefs, and subjective well-being are reported in the present study. In the third session (on average 35.7 days after the first session; $SD = 5.5$ days), longing was re-assessed by means of an adjusted version of the original longing assessment procedure (for details see below). All questionnaires were preceded by detailed written instructions. Approximately one year after data collection, all participants received a letter with feedback on the main results of the study.

Changes in sample size in the second and third sessions. From the original sample, 12 participants did not attend the second session (3 persons had other commitments but returned for the third session, 2 persons dropped out of the study due to illness or new work conditions, and 7 persons refused further participation in the study). This reduced the sample size for control and subjective well-being measures to $N = 287$. Sample attrition in the third session amounted to 16 persons (in addition to the 9 persons who had already permanently dropped out in the second session, 4 persons lacked the time to attend, 1 person became seriously ill, and 2 persons refused any further participation). This reduced the sample size for the third session (re-assessment of longing) to $N = 283$. The group of individuals who missed one or two test sessions did not differ significantly from the group of individuals who completed all three sessions in terms of sex, education, self-rated health, and intelligence. However, there was a trend for participant who dropped out to be younger than the remaining sample (42.1 versus 50.4 years; $t(297) = 2.09$; $p = .04$). This trend appeared to be unrelated to the subject of the present study. Younger participants cancelled test sessions more often because of unexpected work or family obligations.

3.3 Measures

3.3.1 *Measurement of Longing*

Longing was assessed with a mixed idiographic-nomothetic assessment technique, which was successfully used in previous studies to examine motivational constructs (e.g., Brunstein et al., 1998; Emmons, 1989; J. Heckhausen et al., 2001; B. R. Little, 1983; Riediger, 2001; Wrosch & Heckhausen, 2002). Accordingly, the measurement procedure consisted of two phases: a Longing Generation phase (idiographic part) and a Longing Assessment phase (nomothetic part; for an overview, see Table 6). Longing was assessed twice, once in the first session and again in the third session (re-assessment). In addition, an anonymous follow-up checklist of intimate longing contents was completed at the end of the study. The measurement procedure for the assessment of longing had been developed and tested in three pilot studies (see Appendix B). A preliminary version of the procedure described below was used in the third pilot study and proved successful for the elicitation and investigation of longing.

Longing Generation Phase

In the *Longing Generation phase*, participants were asked to self-generate a list of personal longings defined as “strong wishes for persons, things, events, or experiences from your personal past, present, or future that are intensive, enduring, and not easily attainable at present” (for the complete definition given to participants, see Appendix C). This task was supported by a “guided mental journey through life,” in which participants were asked to visualize five different life phases (childhood, youth, young adulthood, middle adulthood, late adulthood) by means of images of significant places, persons, and experiences from each life phase. Depending on the age of the participant, some of these life phases were past, while others were future phases. The experimenter read a text that guided participants in re-experiencing each life phase in as much imaginary detail as possible (see Appendix C). The intention of this procedure was to encourage participants to reflect on their whole life, including its overarching themes and wishes (McAdams, 2001; Staudinger, 2001). After visualizing each life phase, participants were asked to add to their list all longings that were related to this life phase.

After generating a list of idiographic longings, the study definition of longing was read to participants again to remind them that their longings should be emotionally intensive, enduring, and not easily attainable at present. Participants were then asked to (1) exclude the contents that did not qualify as longing according to this definition, and (2) combine contents that were strongly related to one another and thus represented the same longing. Next, they were instructed to rank-order the resulting longings according to their current importance and relevance in their

lives. Subsequently, they were asked to shortly describe their three most important longings. In total, the Longing Generation phase lasted about 45 minutes.

Table 6
Overview of the Procedure for the Assessment of Longing

Phase	Description
1. Longing generation	Generation of a list of personal longings during a guided mental journey through life and ranking according to current importance (Session 1)
2. Longing assessment	Rating of the three most important longings on a questionnaire measuring their structural characteristics, salience, content, controllability, and function (Session 1)
3. Re-assessment of longing	Repeated free generation of longings during guided mental journey and re-assessment of the originally selected three most important longings (Session 3; approximately five weeks after Session 1)
4. Anonymous follow-up checklist of more private longings	Anonymous checklist of more private longings (e.g., own death, infidelity) administered at the end of Session 3

Longing Assessment Phase

In the *Longing Assessment phase*, participants rated their three most important longings on a newly developed questionnaire. The questionnaire consisted of three equal parts, one for each of the three longings. At the beginning of each part, participants were reminded to think of their most (second most, or third most, respectively) important longing when responding to the items.

The present theoretical conceptualization of longing provided the framework for the development of the questionnaire. Accordingly, the questionnaire was designed to assess (1) the six *structural characteristics* of longing (incompleteness, symbolic nature, personal utopia, emotional ambivalence, tritime focus, reflection), (2) the *salience* of longing, (3) the *contents* of longing, (4) longing-specific *control beliefs* (control over the experience and control over the realization of longing), (5) the two *functions* of longing (directionality, managing nonrealizability), and (6) several *additional scales and items* for the characterization of longing (attainability, emotional quality). The scales showed moderate to satisfactory internal consistencies in the third pilot study (see Appendix B). Items that showed extreme skewness and/or kurtosis in the pilot study or increased Cronbach's alpha if deleted were reformulated prior to this study. All items were rated on a six-point scale ranging from 0 (*does not apply at all*) to 5 (*applies very much*) except otherwise indicated.

The order of scales was as follows: Salience scale, content and additional items (Block 1), Structural Elaboration and Longing Function scales (Block 2)⁶, and Longing Control scales (Block 3). Items were presented in random order within Blocks 2 and 3, except for the Tritime Focus scale that, due to the differing response format, was presented at the end of Block 2. In the following, subscales are described in more detail. Item characteristics and intercorrelations are listed in Appendices D and E, respectively.

1. *Scales assessing the structural elaboration of longing.* Six scales were designed to measure the structural elaboration of longings, one for each of the structural characteristics: Incompleteness, Symbolic Nature, Personal Utopia, Emotional Ambivalence, Tritime Focus, and Reflection. Original and translated items are presented in Table 7. With the exception of the Tritime Focus scale, each scale consisted of six statements that were rated in terms of how much they applied to each of the three longings of the person.

The tritime focus was measured with two indicators. First, participants rated how much their longing had to do with their personal past, present, and future (3 items, same 6-point scale format as in the other five structural scales). The tritime focus was indicated by the mean of these three items, with higher values indicating a higher tritime focus. Second, participants distributed 100 points among past, present, and future according to the amount of time they think about these time periods during longing. The variance of these three point sums was used as an index of the tritime focus, with a lower variance indicating a higher tritime focus. For reasons of simplicity, linear transformations were performed so that the index would range from 0 to 5, with higher values indicating a higher tritime focus.

2. *Scale assessing the salience of longing.* The salience of longings was measured with four bipolar items, one each for the frequency, intensity, and duration of longing episodes, and the centrality of the specific longing to the person's self-concept. Thus, this scale measured how salient and central participants perceived each longing in their lives (see Table 7).

3. *Items assessing the content of longing.* Participants rated the degree to which their longings were related to each of 13 life domains on a six-point scale (0 = *has nothing to do with it*; 5 = *has very much to do with it*). The selection of life domains was informed by previous studies on personal goal domains (J. Heckhausen et al., 2001; Nurmi, 1992; Riediger, 2001; Staudinger, Freund, Linden, & Maas, 1999). The 13 life domains were chosen from the four general domains of self (health, physical well-being, self-image/personal characteristics, and spirituality/religious experiences), personal context (financial situation/ possessions, work/education, leisure/hobbies,

⁶ The Structural Elaboration and Longing Function scales were presented in a single block because these two components of longing were only distinguished conceptually when data collection was already in progress.

and living situation/location), interpersonal relationships (partnership, family, and friendships), and society (societal values and politics/world situation). Life domains were presented in a random order. Participants were instructed to be as discriminative as possible in their responding.

Table 7
Overview of Items of the Longing Questionnaire

Subscales	Item Labels	Items	
		English	German
<i>Structural Elaboration</i>			
Incom- pleteness	I1	My longing means that something essential is missing in my life.	Meine Sehnsucht bedeutet, dass in meinem Leben etwas Wesentliches fehlt.
	I2	My longing means that I am missing one of the most important things in my life.	Meine Sehnsucht bedeutet, dass ich eines der wichtigsten Dinge in meinem Leben vermisste.
	I3	As long as my longing is unfulfilled, something essential is missing for me.	Mir fehlt etwas ganz Wesentliches, solange meine Sehnsucht nicht erfüllt ist.
	I4	If my longing were to remain unfulfilled, it would not make a big difference in my life. (R)	Wenn meine Sehnsucht unerfüllt bleibt, fände ich das schade, aber nicht wirklich tragisch. (R)
	I5	If my longing were to be fulfilled, my life would be more complete.	Wenn sich meine Sehnsucht erfüllen würde, wäre mein Leben komplett.
	I6	My life is complete even without that which I am longing for. (R)	Auch ohne das, wonach ich mich sehne, ist mein Leben vollständig. (R)
Symbolic Nature	S1	What I am longing for is heavily filled with meaning.	Das, wonach ich mich sehne, ist mit Bedeutung aufgeladen.
	S2	What I am longing for embodies some higher aim (e.g., success, happiness, or love).	Das, wonach ich mich sehne, verkörpert etwas Großes (z.B. Erfolg, Glück oder Liebe).
	S3	What I am longing for symbolizes something important to me.	Das, wonach ich mich sehne, steht symbolisch für etwas, das mir wichtig ist.
	S4	What I am longing for is important by itself and has nothing to do with some higher aim. (R)	Das, wonach ich mich sehne, ist für sich allein wichtig und nicht, weil es für andere Dinge steht. (R)
	S5	What I am longing for is symbolic for something else.	Das, wonach ich mich sehne, steht symbolisch für etwas Anderes.
	S6	What I am longing for does not hold a deeper meaning. (R)	Das, wonach ich mich sehne, birgt kaum tiefer liegende Bedeutungen. (R)
Personal Utopia	U1	If my longing were fulfilled, it probably would not be as great as it is in my fantasy.	Wenn sich meine Sehnsucht erfüllen würde, wäre es wahrscheinlich nicht so schön, wie ich es mir in meiner Fantasie ausmale.
	U2	Reality will never be the way I long for it to be.	Die Wirklichkeit wird nie so sein, wie ich es mir ersehne.

(table continues)

Table 7. (continued)

Subscales	Item Labels	Items	
		English	German
(Personal Utopia)	U3	I am longing for something too perfect to be true.	Ich sehne mich nach etwas, das eigentlich zu perfekt ist, um Wirklichkeit zu werden.
	U4	My longing is an ideal image of my life.	Meine Sehnsucht ist eine Idealvorstellung von meinem Leben.
	U5	When my longing is fulfilled it will be exactly the way I imagine it. (R)	Wenn sich meine Sehnsucht erfüllt, wird es genauso sein, wie ich es mir vorstelle. (R)
Ambivalent Emotions	U6	My longing is like a dream I am chasing after.	Meine Sehnsucht ist wie ein Traum, dem ich hinterher jage.
	E1	Experiencing my longing is pleasant and unpleasant at the same time.	Meine Sehnsucht zu fühlen, ist beides zugleich: angenehm und unangenehm.
	E2	My longing is a bittersweet feeling.	Meine Sehnsucht ist ein bittersüßes Gefühl.
	E3	Feeling my longing is a bit like enjoying sad music.	Meine Sehnsucht zu fühlen ist ein bisschen wie traurige Musik zu genießen.
	E4	My feeling of longing is both painful and pleasurable.	Mein Sehnsuchtsgefühl ist gleichzeitig schmerzhaft und schön.
	E5	My longing makes me either cheerful or sad, but not both at the same time. (R)	Entweder macht mich meine Sehnsucht froh oder traurig, aber nicht beides zugleich. (R)
Tritime Focus	E6	I experience my longing with mixed feelings.	Ich empfinde meine Sehnsucht mit gemischten Gefühlen.
	T1	My longing has to do with people, things, experiences, or events ... in my past / ... in my present / ... in my future. ^a	Meine Sehnsucht hat zu tun mit Personen, Dingen, Erfahrungen oder Erlebnissen ... aus meiner Vergangenheit / ... aus meiner Gegenwart / ... aus meiner Zukunft. ^a
	T2	When you have this longing, how much do you think about your past, present, or future? Please express the extent of your thoughts in points. You have 100 points. Please distribute these 100 points among the three time periods. Time periods can also be assigned 0 points. ^b	Wenn Sie diese Sehnsucht haben, wie sehr denken Sie dann an Ihre Vergangenheit, Gegenwart oder Zukunft? Bitte drücken Sie den Umfang Ihrer Gedanken in Punkten aus. Sie haben dazu insgesamt 100 Punkte zur Verfügung. Bitte teilen Sie diese 100 Punkte auf die drei Zeitabschnitte auf. Sie können Zeitabschnitten auch 0 Punkte zuordnen. ^b
Reflection	R1	When this longing appears I think for a long time about how far I have gotten in my life.	Wenn diese Sehnsucht auftaucht, denke ich lange darüber nach, wo ich in meinem Leben stehe.
	R2	My longing makes me think a lot about the meaning of my life.	Meine Sehnsucht bringt mich dazu, mir viele Gedanken über den Sinn meines Lebens zu machen.

(table continues)

Table 7. (continued)

Subscales	Item Labels	Items	
		English	German
(Reflection)	R3	My longing often makes me start thinking intensively about myself and my life.	Mein Sehnsuchtsgefühl ist oft der Anfang von einem intensiven Nachdenken über mich und mein Leben.
	R4	When I am having this longing, I think about ways to better shape my life.	Ich denke darüber nach, wie ich mein Leben besser gestalten kann, wenn ich diese Sehnsucht habe.
	R5	When my longing appears, it is just a fleeting feeling, without many thoughts. (R)	Wenn meine Sehnsucht auftaucht, ist das eher ein Gefühl, als dass es mich zum Nachdenken bringt. (R)
	R6	Feeling my longing does not lead me to contemplate over my life. (R)	Das Gefühl meiner Sehnsucht zieht keine ausgedehnten Gedanken über mein Leben nach sich. (R)
<i>Saliency</i>			
Frequency	Sal1	My longing appears ... very rarely / very often.	Meine Sehnsucht tritt ... sehr selten auf / sehr häufig auf.
Duration	Sal2	When my longing appears, it is ... a fleeting feeling / a prolonged feeling.	Wenn meine Sehnsucht auftaucht, ist das ... ein kurzes Gefühl / ein langanhaltendes Gefühl.
Intensity	Sal3	I experience my longing as ... not intense / very intense.	Ich erlebe meine Sehnsucht als ... wenig intensiv / sehr intensiv.
Ego-Centrality	Sal4	My longing is ... unimportant and irrelevant for my life / central to my life.	Meine Sehnsucht ist ... nebensächlich für mein Leben / zentral für mein Leben.
<i>Longing Control</i>			
Control Over Longing Experience	CE1	I can always control my feelings of longing very well.	Ich kann meine Sehnsuchtsgefühle immer sehr gut kontrollieren.
	CE2	Whenever I want to, I have means and ways to quickly distract myself from my feelings of longing.	Wann immer ich will, habe ich Mittel und Wege, um mich schnell von meinem Sehnsuchtsgefühl abzulenken.
	CE3	If my longing becomes too unpleasant, I can immediately change my thoughts to feel better again.	Wenn meine Sehnsucht zu unangenehm wird, kann ich sofort meine Gedanken so ändern, dass ich mich wieder gut fühle.
	CE4	I usually can influence when and in which situations I have this longing.	In der Regel kann ich selbst bestimmen, wann und in welchen Situationen ich diese Sehnsucht habe.
	CE5	I cannot keep my thoughts from constantly revolving around my longing. (R)	Ich kann es nicht verhindern, dass meine Gedanken ständig um diese Sehnsucht kreisen. (R)

(table continues)

Table 7. (continued)

Subscales	Item Labels	Items	
		English	German
(Control Over Longing Experience)	CE6	When I get this longing, I feel completely helpless in face of it. (R)	Wenn mein Sehnsuchtsgefühl hochkommt, fühle ich mich diesem völlig ausgeliefert. (R)
Control Over Longing Realization	CR1	I am sure that I can fulfill my longing sometime.	Ich bin sicher, dass ich mir meine Sehnsucht irgendwann erfüllen kann.
	CR2	I know exactly which steps I should take towards fulfilling my longing.	Ich weiß genau, welche Schritte ich selbst in Richtung meiner Sehnsucht unternehmen kann.
	CR3	Luck will be on my side in the fulfillment of my longing.	Das Glück wird mir bei der Erfüllung meiner Sehnsucht zur Seite stehen.
	CR4	It does not matter what I do, my longing will probably never come true. (R)	Egal was ich tue, meine Sehnsucht wird sich wahrscheinlich nie erfüllen. (R)
	CR5	I myself can do nothing to fulfill my longing. (R)	Ich selbst kann <i>nichts</i> dafür tun, um der Erfüllung meiner Sehnsucht näher zu kommen. (R)
	CR6	I know to whom I should turn to to come a little closer to fulfilling my longing.	Ich weiß, an wen ich mich wenden kann, damit sich meine Sehnsucht ein bisschen mehr erfüllt.
	CR7	Other people cannot do anything for the fulfillment of my longing. (R)	Andere Leute können <i>nichts</i> dafür tun, damit sich meine Sehnsucht erfüllt. (R)
	CR8	When trying to turn my longing into action I usually have no luck. (R)	Wenn ich versuche, meine Sehnsucht in die Tat umzusetzen, habe ich meist kein Glück. (R)
<i>Longing Function</i>			
Directionality	D1	My longing is like a vision towards which I orient my life.	Meine Sehnsucht ist wie eine Vision, nach der ich mein Leben ausrichte.
	D2	My longing shows me clearly what really matters in my life.	Meine Sehnsucht zeigt mir ganz klar, worum es in meinem Leben geht.
	D3	My longing gives a direction to my life.	Meine Sehnsucht gibt meinem Leben eine Richtung.
	D4	My longing makes me realize over and over again what is really important in my life.	Meine Sehnsucht macht mir immer wieder sehr bewusst, was für mein Leben wirklich wichtig ist.
	D5	My longing has no influence on how I live my life. (R)	Meine Lebensgestaltung orientiert sich nicht an meiner Sehnsucht. (R)
	D6	My longing does not direct the future course of my life. (R)	Ich richte mein weiteres Leben nicht nach meiner Sehnsucht aus. (R)
Managing Nonrealizability	MN1	Experiencing my longing partially compensates for something I cannot have in reality.	Meiner Sehnsucht nachzugehen ist ein kleiner Ersatz für etwas, das ich in der Realität nicht haben kann.

(table continues)

Table 7. (continued)

Subscales	Item Labels	Items	
		English	German
(Managing Nonrealizability)	MN2	Experiencing my longing helps me a bit to get over something I do not have any more.	Meine Sehnsucht zu erleben tröstet mich ein bisschen darüber hinweg, dass ich etwas nicht mehr habe.
	MN3	Through my longing I keep my memories of something past alive.	Durch meine Sehnsucht halte ich die Erinnerung an etwas Vergangenes lebendig.
	MN4	Through my longing I can compensate a bit for something missing in my real life.	Ich kann mit meiner Sehnsucht ein bisschen ausgleichen, was mir im wahren Leben fehlt.
	MN5	In my longing I can live out something, which I have to do without in my life at present.	In meiner Sehnsucht kann ich etwas ausleben, auf das ich im wirklichen Leben momentan verzichten muss.
	MN6	My longing helps me to better cope with a difficult experience or loss.	Meine Sehnsucht hilft mir, besser mit einem schwierigen Ereignis oder Verlust umzugehen.
<i>Additional Subscales and Items</i>			
Attainability	--	The fulfillment of my longing is ... unattainable / attainable.	Die Erfüllung meiner Sehnsucht ist... unerreichbar / erreichbar.
Positive Emotions	--	When I have this longing, I feel ... well / happy / alert / fresh / calm / well-balanced / confident / cheerful.	Wenn ich diese Sehnsucht habe, fühle ich mich... wohl / glücklich / wach / frisch / ruhig / ausgeglichen / zuversichtlich / froh.
Negative Emotions	--	When I have this longing, I feel ... unhappy / dissatisfied / sleepy / exhausted / tense / nervous / discouraged / sad.	Wenn ich diese Sehnsucht habe, fühle ich mich ... unglücklich / unzufrieden / schläfrig / ermattet / angespannt / nervös / entmutigt / traurig.

Note. Items with bold labels were included in the final scales, whereas the remaining items were excluded during the process of scale optimization (see Section 4.1.1). (R) = reversed item.

^aThe construct was derived by calculating the mean of the three items.

^bThe construct was derived by calculating the standard deviation of the three values and performing linear transformations so that the final score range was 0 to 5, with higher values indicating a higher tritime focus.

4. *Scales assessing longing-specific control beliefs.* Longing-specific control beliefs were measured with two scales: Control Over Longing Experience (measuring the perceived ability to control the experience of longing; 6 items) and Control Over Longing Realization (measuring the perceived ability to fulfill the longing in actual life; 8 items; see Table 7). The first scale was informed by research on emotion regulation. Emotion regulation refers to processes by which individuals influence which emotions they have, when they have them, and how they experience and express these emotions (Gross, 1998). Accordingly, Control Over Longing Experience measures how much people believe they can influence the onset, course, and ending of longing episodes.

The Control Over Longing Realization scale was designed on the basis of previous research on domain-general control beliefs (e.g., Skinner, 1996). Items were formulated to cover four different aspects of control: general control beliefs, internal agency beliefs, external agency beliefs regarding powerful others, and external agency beliefs regarding chance (two items each).

5. *Scales assessing the function of longing.* The two proposed functions of longing were measured by one scale each: Directionality and Managing Nonrealizability. The Directionality scale measured the extent to which people considered their longing as a vision towards which they directed their lives. Managing Nonrealizability measured whether people regarded their longing as a partial compensation for lost or unrealizable life paths. For both scales, higher scores indicated a higher degree of perceived functionality of longing. Each scale consisted of six items (see Table 7).

6. *Additional subscales and items.* In addition to the scales measuring the structural characteristics, salience, content, function, and controllability of longings, participants rated their longings on several further characteristics. First, a single bipolar item measured the attainability of longings (Table 7). Second, the composition of emotions during longing was assessed in more detail. Using an adjective checklist, participants rated how much they experienced eight positive and eight negative emotions during longing. Positive and negative emotions were thus assessed independently from each other. The checklist comprised 6 emotion pairs from the 12-item short form of the Multidimensional Affect Rating Scale (Steyer, Schwenkmezger, Notz, & Eid, 1997) and two additional emotion pairs (confident-discouraged, cheerful-sad) that were considered relevant to longing. Adjectives were presented in random order and rated on a six-point scale (0 = *not at all*, 5 = *very strongly*). Third, the ratings of 13 life domains of longing described before were used to derive an index of the extension of the longing across different life domains. This index counted the number of life domains rated as very relevant (4 or 5 on the response scale).

Re-Assessment of Longing

In the third test session, participants were again asked to generate a list of personal longings according to the procedure used in the first session (re-assessment of the free recall of longing contents). At the beginning of the session, participants were instructed that the thoughts and longings that would come into their mind today could be very similar to the ones in the first session, or very different, and that either would be just fine.

Following this Longing Generation phase, they received a copy of their three most important longings selected in the first session and indicated whether each of these original longings was also part of the newly generated list of longings. This provided an index of the temporal stability of longing contents as judged by the participants. In some cases, participants

failed to provide this information although some of their longing descriptions from the third session clearly matched the ones from the first session. In these cases, longing contents were considered stable. After matching their two longing lists, participants re-assessed their three original longings on the Longing Questionnaire (re-assessment of the structural elaboration, salience, content, controllability, and function of longing).

Anonymous Follow-Up Checklist of More Private Longings

Longing representations may include wishes that are considered as socially undesirable or very intimate (e.g., criminal activities or sexual experiences) so that they are not readily disclosed in the setting of a research study. To nevertheless assess the frequency of these “more private” longing contents, an anonymous follow-up checklist was developed and administered at the end of the third test session. The purpose of this follow-up checklist was to increase participants’ willingness for self-disclosure. Participants completed the checklist separately from the rest of the questionnaires, put it into a sealed envelope, and left it in a box just before leaving the test room. Participants were informed that this checklist did not contain an identification number and thus could not be linked to any personal information or to the rest of their data.

The checklist first asked participants to indicate whether they had any longings they did not report during the course of the study because they found them too intimate to reveal. In the case of a “yes”-response, they were asked to mark one or more content categories for these private longings. Content categories included sexual experiences, infidelity, own death, death of others, revenge, dominating others, possessions of other people, physically or mentally hurting others, other kinds (with a line provided for their description), and a category “I don’t want to give this information”. When participants initially denied the presence of any private longings, but checked off one or more specific content categories, their initial response was recoded as a “yes”-response (this happened in 20 cases or 7.3 %).

3.3.2 Measurement of Subjective Well-Being

Four well-established measures were chosen for the assessment of subjective well-being (see Table 8). Three measures were included for the assessment of happiness (positive and negative affect, global life satisfaction, domain-specific life satisfaction), and one measure was included for the assessment of positive psychological functioning. Since self-report measures were used for the assessment of well-being in this study, items were carefully screened for content overlap with items of the Longing Questionnaire. No content overlap was found. For all subjective well-being scales, the response format was held consistent with the Longing Questionnaire, that is, participants responded on six-point scales ranging from 0 to 5 (with

varying anchors; for details see below). Descriptive information for all indicators of subjective well-being can be found in Appendix F.

Table 8
Overview of Scales for the Measurement of Subjective Well-Being

Constructs	Scales and Modifications
<i>Happiness</i>	
Positive and negative affect	Multidimensional Affect Rating Scale (Multidimensionaler Befindlichkeitsfragebogen; Steyer et al., 1997); modification of instruction and addition of four items
Life satisfaction (global)	Satisfaction With Life Scale (SWLS; Diener, Emmons, Larsen, & Griffin, 1985)
Life satisfaction (domain-specific)	Pressure-to-Change Scale (Veränderungsdruckskala; Filipp & Ferring, 1991); modification of life domains
<i>Positive Psychological Functioning</i>	
Positive psychological functioning	Ryff Inventory (Ryff, 1989)

Note. For each questionnaire, the response format was held consistent with that of the Longing Questionnaire (six-point scale ranging from 0 to 5; for scale anchors, see text).

Positive and Negative Affect

Long-term positive and negative affect was measured with the Multidimensional Affect Rating Scale developed by Steyer et al. (1997). The Multidimensional Affect Rating Scale was originally designed to measure three bipolar dimensions of well-being with four adjective pairs each. The dimensions are positive mood – negative mood (measured by happy-unhappy, well-unwell, satisfied-dissatisfied, good-bad), alertness – fatigue (measured by rested-exhausted, fresh-tired, alert-sleepy, awake-weak), and calmness – restlessness (measured by relaxed-tense, composed-nervous, calm-restless, and well-balanced -anxious). Two adjective pairs were added to make the scale comparable with the emotion checklist used to assess longings: cheerful-sad and confident-discouraged. The time frame of instruction was also modified. Instead of evaluating their current mood, participants indicated on a six-point scale how frequently they had experienced each emotion during the past year (0 = *very infrequently*, 5 = *very frequently*). The valence aspect of emotional well-being was most relevant for the present study. Therefore, all adjectives with positive valence were aggregated into a Positive Affect scale and all adjective with negative valence were aggregated into a Negative Affect scale. Internal consistencies of these two scales were good; .94 for Positive Affect and .93 for Negative Affect.

Global Life Satisfaction

Participants completed the 5-item Satisfaction With Life Scale (SWLS; Diener et al., 1985) that assesses overall, domain-general life satisfaction. This scale contains statements such as “In most ways, my life is close to my ideal” or “If I could live my life over, I would change almost nothing”, which participants rated on a six-point scale (0 = *does not apply at all*, 5 = *applies very much*). Previous studies have demonstrated the high test-retest stability and other desirable psychometric characteristics of the SWLS (Pavot & Diener, 1993). In the present study, the internal consistency was .85.

Domain-Specific Life Satisfaction

According to Schwarz and Strack (1985), domain-general indicators of life satisfaction are confounded with current mood and context, as well as social desirability. In contrast, domain-specific life satisfaction items are less influenced by such confounds. Therefore, a modified version of the Pressure-to-Change Scale (Veränderungsdruckskala; Filipp & Ferring, 1991) was administered as a domain-specific indicator of (lack of) life satisfaction. Using a six-point scale (0 = *not at all*, 5 = *very much*), participants rated the extent to which they desired a change in 13 domains of life: (1) health, (2) physical well-being, (3) self-image/ personal characteristics, (4) spirituality/ religion, (5) financial situation / possession, (6) work/education, (7) leisure/hobbies, (8) living situation/location, (9) partnership, (10) family, (11) friendships, (12) societal values, and (13) politics/world situation. Life domains only partly overlapped with the original instrument by Filipp and Ferring (1991) and were the same as those used to assess life domains of longings. They represent areas that are central to the lives of healthy younger and older adults as reflected in their goals (J. Heckhausen, 1997; Nurmi, 1992; Riediger, 2001). Scores for the 13 life domains were aggregated to form an index of the average domain-specific desire for change (or lack of life satisfaction). The internal consistency was .79.

Positive Psychological Functioning

The Ryff Inventory (Ryff, 1989) was used as a measure of positive psychological functioning. According to Ryff (1989) and Ryff and Keyes (1995), common measures of subjective well-being have little theoretical grounding and therefore neglect important aspects of psychological functioning emphasized in theories of psychological health and well-being. In response to this criticism, the Ryff Inventory was developed to assess six dimensions of positive psychological functioning often mentioned in the writings of humanistic theorists. These include (1) self-acceptance (a positive attitude toward oneself and one’s past and a differentiated view of the multiple aspects of the self), (2) positive social relations (warm, satisfying, trusting

relationships with others and a capacity for love and empathy), (3) autonomy (self-determination and independence from social pressures as well as regulation of thought and behavior from within), (4) environmental mastery (sense of mastery and competence in managing the environment, effective use of surrounding opportunities, and participation in activities outside of the self), (5) purpose in life (having goals and a sense of directedness, experiencing a sense of purpose in and meaning to life), and (6) personal growth (a feeling of continued development and a wish to change in ways that reflect more self-knowledge and directedness). In the present study, each subscale comprised nine items, amounting to a total of 54 items. Items were rated on a six-point scale (0 = *does not apply at all*, 5 = *applies very much*). Ryff (1989) reported evidence for the subscales' reliability and validity. In the present study, the internal consistency coefficients were as follows: Self-Acceptance, .86; Positive Social Relations, .79; Autonomy, .64; Environmental Mastery, .84; Purpose In Life, .72; and Personal Growth, .66.

Aggregation of Subjective-Well-Being Scales

For reasons of parsimony, subjective well-being scales were aggregated for analyses. In line with the literature on subjective well-being (e.g., McGregor & Little, 1998; Ryan & Deci, 2001; Ryff & Keyes, 1995), positive affect, (absence of) negative affect, global, and domain-specific life satisfaction were aggregated into an index of happiness, and the six subscales of the Ryff Inventory were aggregated into an index of positive psychological functioning.

The four scales combined into happiness were all moderately to highly interrelated, which replicates McGregor and Little's (1998) factor analysis of subjective-well-being measures. Absolute correlations ranged from .46 for Positive Affect and Pressure-to-Change to .66 for Positive and Negative Affect (all $ps = .001$). The Negative Affect and Pressure-to-Change scales were reversed prior to aggregation. The internal consistency of Happiness was .85.

The six subscales of the Ryff Inventory were also moderately to highly interrelated, which replicates previous studies (e.g., Ryff & Keyes, 1995). Intercorrelations ranged from .24 for Autonomy and Positive Social Relations to .74 for Environmental Mastery and Self-Acceptance (all $ps = .001$). The internal consistency was .83.

3.3.3 Additional Measures

Three additional scales were included as control variables. These measured socially desirable responding and two aspects of domain-general control beliefs (general self-efficacy and general emotion regulation).

Socially Desirable Responding

The Social-Desirability-Scale 17 (SES-17; Stöber, 1999) was used to assess the tendency for socially desirable responding. Using a two-point scale (1 = *yes*, 0 = *no*), participants indicated their agreement with 16 items⁷ describing either socially desirable behaviors that are very unlikely to occur (e.g., “I always admit my mistakes openly and face the potential negative consequences.”) or socially undesirable behaviors that are very likely to occur (e.g., “Sometimes I only help because I expect something in return.”). The internal consistency was .75.

General Self-Efficacy

Participants completed the General Self-Efficacy Scale (Schwarzer & Jerusalem, 1995) as a global measure of personal action control and agency. This scale consisted of 10 items (e.g., “I can always manage to solve difficult problems if I try hard enough.”; “If I am in trouble, I can usually think of a solution.”). The response format was adapted to the format of the remaining questionnaires (6-point scale; 0 = *does not apply at all*, 5 = *applies very much*). The internal consistency was .88.

General Emotion Regulation

Due to a lack of well-established measures, 10 new items were formulated to assess general emotion regulation abilities. Item formulation was guided by the literature on emotion regulation (Gross, 1998, 1999). The items largely paralleled the items of the Control Over Longing Experience scale. They included two items for general emotion control (e.g., “I can always control my feelings very well.”), and four items each for the ability to up-regulate positive and the ability to down-regulate negative emotions (e.g., “When I have unpleasant feelings, I feel totally helpless in face of them.”, “When I want to feel good, I can direct my thoughts to the good aspects of the situation.”). Items are presented in Appendix G. The response format was equal to the General Self-Efficacy Scale.

This newly developed scale had a satisfactory reliability as well as (preliminary) convergent and discriminant validity. The internal consistency was .80. Correlations were .56 with general self-efficacy, .63 with happiness, and .59 with positive psychological functioning (all $ps = .001$; convergent validity). Correlations with intelligence measures were nonsignificant; with mental speed, $-.03$; with verbal knowledge, $.06$ (both $ps > .05$; divergent validity). In addition, there were small positive correlations with age, $.18$; and socially desirable responding, $.28$ (both $ps = .001$).

⁷ One item of the SES-17 scale was excluded due to its low applicability to the present sample: “I have tried illegal drugs (for example, medication, marijuana, or something “similar”) in the past.”

3.4 Statistical Analyses

In a first step, the Longing Questionnaire was analyzed and optimized using a series of confirmatory factor analyses (CFA). Only ratings of each person's most important longing were analyzed at this step. A confirmatory rather than an exploratory approach was chosen for scale development because in the present study, items were formulated in a top-down, theory-guided way. Starting with the construct of longing, different facets and dimensions were identified based on the literature. For each hypothesized dimension, a set of items was created. In such a context when the focus is on validating a model that has been specified *a priori*, confirmatory approaches are most appropriate (Hertzog, 1989; Nesselrode & Baltes, 1984). Moreover, exploratory analyses can lead to interpretive ambiguity if the factors are correlated, which was to be expected in the present study. Subsequently, using multiple-group and repeated-measurement models, the consistency of the optimized scales' measurement structure was tested across (1) the three longings of each person, (2) three age groups, and (3) baseline and retest assessments. As a result of these initial analyses, final longing scales were derived and outliers were adjusted on the level of aggregated items.

The second and third steps included the testing of general and age-related hypotheses about longing characteristics employing uni- and multivariate analyses. Fourth, hierarchical regression analyses were performed to examine the relationship between longing characteristics and the hypothesized outcomes (i.e., the functions of longing as a proximal, specific outcome and subjective well-being as a distal, general outcome). Fifth and finally, meaningful subgroups with different longing profiles were derived by means of cluster analysis of the six Structural Elaboration scales. Hierarchical regression analyses were used to test whether cluster membership moderated previously found associations of longing characteristics with age and subjective well-being. In the following, procedures of data preparation are reported and statistical analyses are described in detail.

3.4.1 Data Preparation

Treatment of Missing Data

Missing data were not replaced in the present study. Although various methods have been proposed for substitution of missing values including mean imputation (the arithmetic mean is substituted for the missing value) and regression imputation (a regression equation is generated on the basis of complete cases that is subsequently used for the estimation of missing values), these methods can lead to biased data (Byrne, 2001; Wothke, 2000). For example, variances are reduced, frequency distributions are changed, and covariances are inflated as a result of data

imputation. Furthermore, in the confirmatory factor models, missing data did not constitute a problem because available Structural Equation Modeling (SEM) software programs such as AMOS (Analysis of Moment Structures; Arbuckle, 2003) allow the estimation of parameters from incomplete data (for details, see below).

Treatment of Outliers

Univariate normality and outliers. Distributions of all variables were screened for deviation from normality and outliers. Due to the relatively large sample size, conventional significance tests for skewness and kurtosis were considered overly sensitive and were thus not conducted (see Tabachnik & Fidell, 1996). Inspection of the absolute skewness and kurtosis values revealed that deviation from zero was within reasonable limits and did not exceed the suggested cut-off values of 3.0 for skewness and 10.0 for kurtosis (Kline, 1998). Furthermore, since the impact of departure from zero skewness and kurtosis on statistical analyses diminishes in large samples (Tabachnik & Fidell, 1996), non-transformed variables were used in all analyses.

The treatment of univariate outliers differed somewhat for variables from the Longing Questionnaire and the remaining variables. For data from the Longing Questionnaire, outliers were first adjusted at the level of aggregated items (i.e., aggregation across the three longings), and then at the level of subscales. For the remaining variables, outliers were only adjusted at the level of subscales. Scores were considered outliers if they were more than three standard deviations above or below the mean. Outliers were assigned the raw score corresponding to a z -score of ± 3 . Descriptive statistics (mean, standard deviation, minimum, maximum, skewness, kurtosis, number of missing values, and number of outliers) for all central variables used in the present analyses can be found in Appendices D (for item level) and F (for scale level).

Multivariate outliers. Multivariate outliers were inspected in central CFAs and in all regression analyses. Cases were considered multivariate outliers if their Mahalanobis distance was significant at $p < .001$ (Tabachnik & Fidell, 1996). In central CFAs, when multivariate outliers were detected, analyses were repeated without these cases and are reported in a footnote. When multivariate outliers were detected in regression analyses, outlying cases were deleted prior to analysis.

3.4.2 Optimization of the Longing Questionnaire With Confirmatory Factor Analysis

Method of Estimation

All CFA analyses were performed with AMOS 5 (Arbuckle, 2003). In order to save variance and covariance information, the covariance matrix was chosen for analysis. The mean vector was also analyzed, as this is required for the analysis of incomplete data. Thus, all models

were mean-and-covariance-structure models. Model parameters were estimated with the Full Information Maximum Likelihood (FIML) algorithm, which allows the estimation of parameters based on incomplete data. FIML assumes that data are missing at random (i.e., that the cause of the missingness is due to factors independent from the variables with missing data) and multivariate normally distributed. The FIML method does not impute missing data, but uses all data observed for each case to compute maximum likelihood estimates of parameters. Each distinct pattern of missing data is treated as a separate group in a multiple-group structural equation model. Parameters (and their standard errors) are then estimated by setting equality constraints for variances, covariances, means, and intercepts across the different patterns of missing data. In simulation studies, the FIML algorithm outperformed most common methods of handling missing data, including data deletion and substitution (Enders & Bandalos, 2001; Wothke, 2000).

Evaluation of Overall Model Fit

After checking for the adequacy of parameter estimates, the overall model fit was evaluated on the basis of three goodness-of-fit indices. The choice of reported fit indices followed recommendations by Hu and Bentler (1998, 1999). Fit indices included (1) the χ^2 -value with its associated degrees of freedom and probability level, (2) the Root Mean Square Area of Approximation (RMSEA) with associated 90% confidence interval, and (3) the Comparative Fit Index (CFI).⁸

The χ^2 -value represents the absolute magnitude of discrepancy between the empirical and the reproduced covariance matrices, with lower values indicating better fit. It is, however, overly sensitive to sample size (i.e., it easily rejects models in large samples) and to the violation of multivariate normality (departures from normality increase the χ^2 -value; Klem, 2000).

The RMSEA represents the average magnitude of discrepancy between the elements in the empirical and reproduced covariance matrices, expressed per degree of freedom (thus taking into account the parsimony of the model). It can range from 0.0 to 1.0, with lower values indicating better fit. According to Brown and Cudeck (1993), RMSEA values below .05 indicate close fit, values between .05 and .08 indicate fair fit, and values greater than .10 indicate poor fit. Hu and Bentler (1999) suggest a value of .06 to be indicative of good fit if the sample size is reasonably large. Because mean-and-covariance-structure models typically exhibit worse fit than

⁸ Hu and Bentler (1998) also recommend reporting the standardized root-mean-square residual (SRMR). However, this index is not computed with the Full Information Maximum Likelihood algorithm in AMOS 5.

covariance models (Aiken, Stein, & Bentler, 1994; Pomplun & Omar, 2003), in this study, values of .06 were considered as good fit, and values of .08 were considered as acceptable fit.

The *CFI* measures the incremental fit of the hypothesized model over a null model, in which all observed variables are uncorrelated and means are freely estimated.⁹ The index is adjusted for sample size. The CFI can range from 0.0 to 1.0, with higher values indicating closer fit. A cut-off value close to .95 has been suggested (Hu & Bentler, 1999). To account for the decrease in CFI that must be expected in mean-and-covariance-structure models (see above), values of .90 and .95 were operationally defined as acceptable and good fit, respectively (see also Aiken et al., 1994; Pomplun & Omar, 2003).

Testing for Parameter Invariance in Multiple-Group and Repeated-Measurement Analyses

To test the consistency of the derived longing subscales across the three longings of each person, three age groups, and baseline and retest assessments, multiple-group or repeated-measurement models were specified. In repeated-measurement models, all latent factors were specified as intercorrelated, as were corresponding residual variances of indicators. For example, covariances were allowed between the residual variances of the same item rated at baseline and re-assessment (or rated for Longings 1, 2, and 3, respectively). Residual variances do contain systematic unique variance of indicators, which should affect responses at all measurement occasions.

Parameter invariance across multiple groups (age groups) or measurements (three longings, baseline and retest) was tested by successively imposing equality constraints on corresponding parameters and evaluating the change in overall model fit. Two types of parameters can be distinguished in structural equation models: parameters at the *measurement level* (factor loadings, intercepts, and residual variances of indicators) and parameters at the *structural level* (factor variances, covariances, and means; Little, 1997). To first establish construct comparability (or measurement equivalence) across groups or measurement occasions, testing for model invariance always started with constraining to be equal the factor loadings of indicators, followed by the intercepts of indicators. Following Byrne (2001) and Little (1997), the residual variances of indicators were not constrained to be invariant since they contain both unique and error variance. Assuming that measurement error is random, it makes little sense to expect it to be invariant across groups or measurement occasions.

⁹ In the 4.0 version of AMOS, the default comparison model used to determine relative fit indices such as the CFI constrained all means to zero. As means are mostly different from zero, this specification resulted in considerably inflated relative fit indices. Therefore, the comparison model was adjusted in the AMOS 5 program.

After equivalence at the measurement level (factor loadings and intercepts) was ascertained, the invariance of structural parameters (factor variances, covariances, and means) was successively tested to examine between- or within-group differences. Whenever a set of parameters appeared to be invariant across multiple groups or measurements (as indicated by a nonsignificant decrease in fit compared to the preceding model), equality constraints were maintained when testing the next set of parameters.

As suggested by Little (1997), two different strategies were used to decide about parameter invariance: a statistical and a modeling rationale. The *statistical rationale* tests the difference in X^2 between two models (with degrees of freedom equal to the difference in their degrees of freedom). If the test is nonsignificant, the models are considered comparable. However, in large samples and complex models with numerous constrained parameters, the X^2 -test is overly sensitive and easily rejects the more restricted model. In contrast, a *modeling rationale* is less conservative as it considers the overall fit of the more restricted model using multiple fit indices. If a model with numerous constraints evinces an adequate overall fit (indicated by acceptable RMSEA- and CFI-values), then the set of constraints are considered a reasonable approximation of the data.

Following Little (1997), a modeling rationale was used for the measurement level (factor loadings and intercepts of indicators) and a statistical rationale for the structural level (factor variances, covariances, and means). At the measurement level, numerous sources of misfit, both random and systematic, must be expected. As long as the overall model fit is acceptable, such lack in empirical accuracy does not diminish the theoretical meaningfulness of the more restricted model. In contrast, the structural level reflects interpretable, error-free effects between constructs. Thus, the more precise statistical rationale is appropriate.

3.4.3 *Derivation of Longing Subgroups With Cluster Analysis*

Longing subgroups were derived by cluster analysis of the six Structural Elaboration scales (Incompleteness, Symbolic Nature, Personal Utopia, Ambivalent Emotions, Tritime Focus, and Reflection) aggregated across the three longings. The goal of cluster analysis is to divide a group of individuals into subgroups (or clusters) by minimizing differences within clusters and maximizing differences between clusters. In this study, similarities between individuals were quantified by Euclidian distance between raw scores. A two-step clustering procedure was used that combined a hierarchical with a non-hierarchical cluster algorithm. First, Ward's hierarchical clustering procedure was applied. Step by step, this procedure combines the two clusters whose fusion results in the minimum increase in the sum of squared distances between cluster centers, starting with the most fine-grained partition (i.e., each cluster containing one person). The ideal

number of clusters was determined using multiple criteria as suggested by Milligan and Cooper (1985): (1) an atypical decrease in between-cluster variance and an atypical increase in within-cluster variance, (2) a maximal difference between the Pseudo- F statistic (indicating the separation of all clusters) and the Pseudo- t^2 statistic (indicating the separation of the two clusters joined at the current step), and (3) a peak in Sarle's cubic clustering criterion that takes into account the variance accounted for by the clusters and the dimensionality of the between-cluster variation.

Since Ward's procedure does not allow the reallocation of persons who may have been poorly classified at an early stage of the analysis, the derived cluster solutions were subsequently used as initial cluster centers for the non-hierarchical K -Means clustering procedure. In this procedure, each individual is reallocated to the most similar cluster based on his Euclidian distance, and cluster centers are recomputed. This procedure is repeated until the largest change in any cluster center is less than 1 %, or after a maximum of 10 iterations.

To ensure the validity and stability of cluster solutions, the clusters were tested for replicability across two random halves of the sample (Breckenridge, 1989). In each half, the two-step clustering procedure (Ward followed by K -Means) was applied and the cluster solutions of both halves were compared for agreement. Participants of each half were assigned to the most similar cluster of the other half based on their Euclidian distance. Cohen's Kappa was calculated to evaluate the agreement of the new cluster assignment with the original one, and the two resulting κ -values were averaged. Since κ -values can vary considerably between random splits of the sample, this procedure was applied ten times, and the median of the ten κ -values was determined (Asendorpf, Borkenau, Ostendorf, & Van Aken, 2001). A median κ -value of at least .60 is usually considered acceptable.

3.4.4 Adjustment of the Level of Significance

For all analyses except latent analyses (CFAs), the alpha-level was adjusted to $p < .01$ to account for effects of multiple testing. Results significant at $p < .05$ were considered statistical trends and are reported if they were in support of theoretical predictions. These criteria are less conservative than the often-used Bonferroni-adjustment procedure that divides .05 by the number of repeated analyses. Although the Bonferroni-adjustment clearly reduces the likelihood of finding significant effects that do not really exist, it also eliminates the chance of finding smaller effects (Cohen, 1990). Given the scarcity of empirical information on longing, it appeared

therefore justified to apply a less conservative criterion for adjusting the level of significance in the present study.¹⁰

¹⁰ Significance levels for CFAs (e.g., overall model fit) were not adjusted, because in these cases, the goal was to accept the null hypothesis (i.e., no misfit between the empirical and reproduced covariance matrix, or between two nested models).