APPENDICES

TABLE OF CONTENTS

APPENDIX A THEORETICAL BACKGROUND ... 196

APPENDIX B MEASURES ... 198

B. 1 Profile-Defining Psychological Measures ... 198

B. 1. 1 Cognitive Functioning ... 198

B. 1. 2 Personality and Self-Related Functioning ... 214

B. 1. 3 Social Integration ... 216

B. 2 Cross-Disciplinary Measures ... 217

B. 3 Reliabilities Over Time for Measures of Personality, Self-Related Functioning, and Social Integration ... 218

APPENDIX C RESULTS ... 219

C. 1 Identification and Description of Heterogeneity and Its Correlates: Subgroup Differences in Psychological Profiles and External Correlates at Baseline ... 219

C. 1. 1 Results of Cluster Analysis: Comparison to Previous Systemic-Wholistic Work From the Berlin Aging Study ... 219

C. 1. 2 Subgroups Extracted From the 6-Year Longitudinal BASE Sample: Profile Characteristics on Psychological Measures at Baseline Assessment ... 220

C. 1. 3 Subgroups Extracted From the 6-Year Longitudinal BASE Sample: Profile Characteristics on Cross-Disciplinary Measures at Baseline Assessment ... 222

C. 1. 4 Checking Assumptions of Statistical Tests ... 223

C. 2 Differential Development Over Time and Its Underlying Mechanisms: Subgroup Differences in Change of Psychological Profiles Over Six Years and the Role of External Correlates ... 227

C. 2. 1 Stability and Change at the Level of the Subgroup-Defining Measures ... 227

C. 2. 1. 1 Schaie's Procedure to Adjust for Error of Measurement ... 227

C. 2. 1. 2 Multilevel Growth Curve Modeling ... 230

C. 2. 2 The Role of External Correlates For Subgroup Change Over Time ... 240

C. 3 Outcomes of Heterogeneity and Differential Development Over Time: Consequences of Subgroup Differences in Psychological Profiles and Change Over Six Years ... 245

APPENDIX D DISCUSSION ... 246

D. 1 Gender Differences in the Total Cross-Sectional and the 6-Year Longitudinal BASE Sample ... 246

APPENDIX A THEORETICAL BACKGROUND

Previous Research on Differential Development in Old Age from a Person-Oriented Perspective. Table A.1 provides examples of person-oriented studies in the context of research on heterogeneity and differential development. These examples of previous work have been organized by several criteria: (a) the dimensions used to define the subgroups, (b) whether or not a given study explicitly included participants in the Fourth Age, for which age 85 years can be regarded a heuristic threshold, (c) the number of measurement occasions (i.e., waves), and (d) the sample size. The strategy to define the subgroups (top-down vs. bottom-up) is also indicated.

Table A. 1
Examples of Research Applying a Person-Oriented Approach to the Study of Heterogeneity and Differential Development in Old Age

Projects	Group-Defining Dimensions	85+ Years[1]	Waves	*N*
Betula Study (Lövdèn, Bergman et al., in press)[2]	Cognition	X	3	500
New England Centenarian Study (Evert et al., 2003)[2]	Health	X	1	424
Duke Longitudinal Study of Aging (Manton et al., 1986)[2]	Cognition	–	11	267
Seattle Longitudinal Study (Schaie, 1990)[2]	Cognition	–	5	628
National Long-Term Care Survey (Manton & Land, 2000)[3]	Functional Health, Mortality	X	4	20,000
Epidemiological Survey in Canberra (Jorm et al., 1998)[3]	Health, Cognition	X	1	977
Wisconsin Longitudinal Study (B. Singer et al., 1998)[3]	Depression, Well-Being	–	3	1,172
Australian Longitudinal Study of Aging (Andrews et al., 2002)[3]	Functional Health, Cognition	–	1	1,043
MacArthur Studies on Successful Aging (Berkman et al., 1993)[3]	Functional Health, Cognition	–	1	1,354
Normative Aging Study (Aldwin et al., 2001)[2]	Health, Personality, Mortality	–	4	1,515
Seattle Longitudinal Study (Bosworth & Schaie, 1997)[2]	Health, Social Integration	–	1	387
Kansas City Study of Adult Life (Neugarten et al., 1968)[2]	Personality, Activity, Well-Being	–	1	59
Americans' Changing Lives Survey (Garfein & Herzog, 1995)[3]	Functional Health, Cognition, Well-Being, Productivity	X	1	1,644
OCTO Study (Zarit et al., 1993)[3]	Functional Health, Cognition, Sensory Functioning	X	1	320
Berlin Aging Study (Smith & Baltes, 1997)[2]	Cognition, Self & Personality, Social Integration	X	1	510
Berlin Aging Study (Smith & Baltes, 1998)[2]	Health, Cognition, Self & Personality, Well-Being, Social Integration, SES	X	1	508
H-70 Study, Gothenburg (Maxson et al., 1996)[2]	Health, Cognition, Well-Being, Social Contacts	–	3	335

Note. [1] This category indicates whether or not the study explicitly focused on participants older than age 85 (X = yes). [2] Bottom-up strategy to subgroup identification. [3] Top-down strategy to subgroup identification.

APPENDIX B MEASURES

B. 1 Profile-Defining Psychological Measures

B. 1. 1 Cognitive Functioning

Perceptual Speed: Digit Letter

BaseCog Dokumentation: Zahlen und Buchstaben 1 14.6.1990 – 1 –

64. Zahlen-Buchstaben-Test

(Die Tafel mit den Beispielaufgaben vorlegen)

Jetzt habe ich ein Zahlen-Buchstaben-Spiel.
Bei diesem Spiel gehören bestimmte Zahlen und Buchstaben immer zusammen. Hier oben (*hindeuten*) **sehen Sie die Zahlen 1 bis 9. Zu jeder Zahl gehört ein bestimmter Buchstabe. Das K gehört zur 1, das P zur 2 usw.**

Hier (*auf Beispiele deuten*) **sehen Sie nur die Zahlen, die Buchstaben fehlen. Sie sollen einfach die Buchstaben, die zu den Zahlen gehören, nennen. Wenn hier also die 2 steht, so sehen Sie nach oben zur 2 und dem Buchstaben, der zur 2 gehört. Diesen Buchstaben - hier also das P - sollen Sie nennen. Welcher Buchstabe gehört dann zur 1?**

(ST bearbeitet Beispiel 1 mit Hilfe und Rückmeldung vom Interviewer. Beispiel 2 soll ST allein lösen. Falls dies nicht gelingt, mit Hilfe von FTA Beispiel 1 wiederholen. Dann ST noch einmal an Beispiel 2 arbeiten lassen)

Bitte arbeiten Sie jetzt SO SCHNELL WIE MÖGLICH weiter. Bitte sagen Sie mir immer, welcher Buchstabe zu welcher Zahl gehört. Wenn Sie mit einer Reihe fertig sind, schlage ich die Seite für Sie um. Fangen Sie jetzt bitte an.

(Stoppuhr! Bearbeitungszeit 3 Min.; die Tafeln nacheinander umklappen. Im Abstand von 1 Min. markieren, an welcher Stelle der Aufgabe - Tafelnummer, Itemnummer - sich ST befindet.)

Erreichter Wert:

	Tafel	Item
nach 60 s	___	___
nach 120 s	___	___
nach 180 s	___	___

verweigert = 7,97,997,9997,99997; weiß nicht = 8,98,998,9998,99998; trifft nicht zu = 9,99,999,9999,99999

Zahlen und Buchstaben I

1	2	3	4	5	6	7	8	9
K	P	U	F	M	C	S	H	A

Beispiel 1:

2	1	3	2
P	?	?	?

BaseCog Dokumentation: Zahlen und Buchstaben 1

1	2	3	4	5	6	7	8	9
K	P	U	F	M	C	S	H	A

Beispiel 2:

2	1	3	2
?	?	?	?

BaseCog Dokumentation: Zahlen und Buchstaben 1

1	2	3	4	5	6	7	8	9
K	P	U	F	M	C	S	H	A

1.

4	1	2	7	9	5
?	?	?	?	?	?

BaseCog Dokumentation: Zahlen und Buchstaben 1

Zahlen und Buchstaben II

1	2	3	4	5	6	7	8	9
L	R	O	D	N	B	T	G	E

Beispiel 1:

2	1	3	2
R	?	?	?

BaseCog Dokumentation: Zahlen und Buchstaben 2

1	2	3	4	5	6	7	8	9
L	R	O	D	N	B	T	G	E

Beispiel 2:

BaseCog Dokumentation: Zahlen und Buchstaben 2

2	1	3	2
?	?	?	?

1	2	3	4	5	6	7	8	9
K	P	U	F	M	C	S	H	A

BaseCog Dokumentation: Zahlen und Buchstaben 2

1.

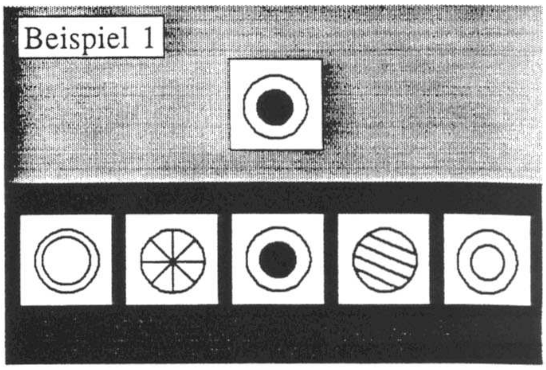
4	1	2	7	9	5
?	?	?	?	?	?

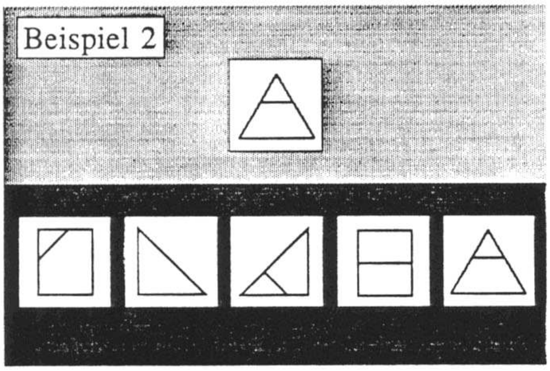
Perceptual Speed: Identical Pictures

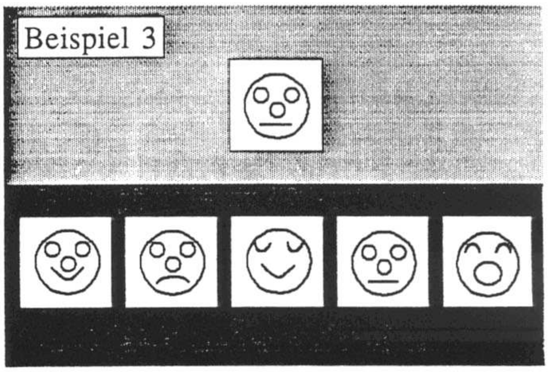
BaseCog Dokumentation: Gleiche Bilder 14.6.1990 – 1 –

ERKLÄRUNG

Bei den nächsten Aufgaben zeigen wir Ihnen oben immer ein Bild. Unter den fünf Bildern unten sollen Sie SO SCHNELL WIE MÖGLICH dasjenige heraussuchen und berühren, das genauso aussieht wie das obere.

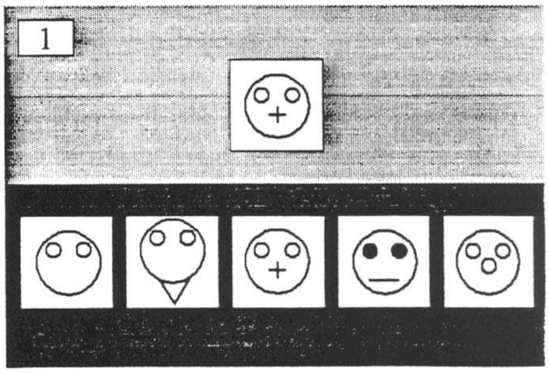


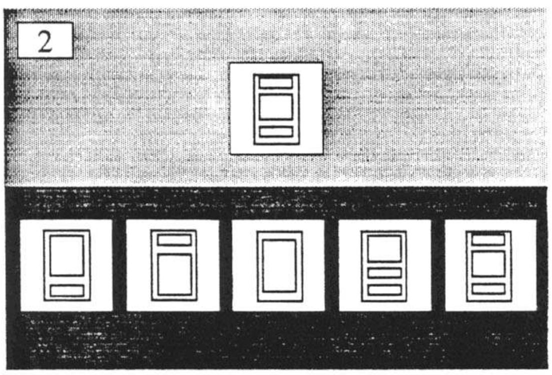


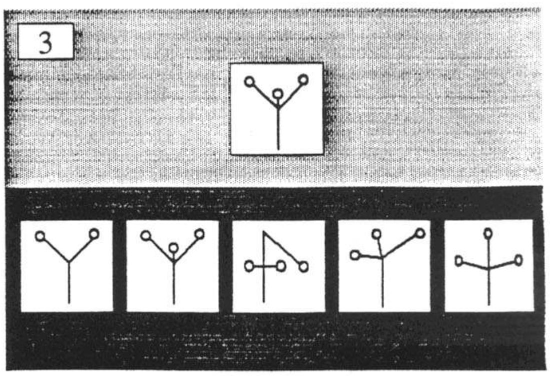


Jetzt kommen weitere Aufgaben in dieser Art. Bitte finden Sie das Bild SO SCHNELL WIE MÖGLICH. Mit dem Berühren von START können Sie beginnen.

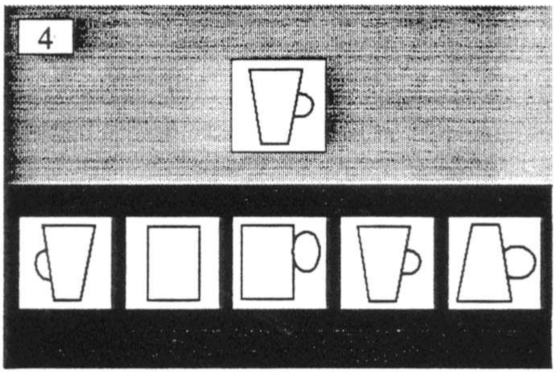
START ⇨

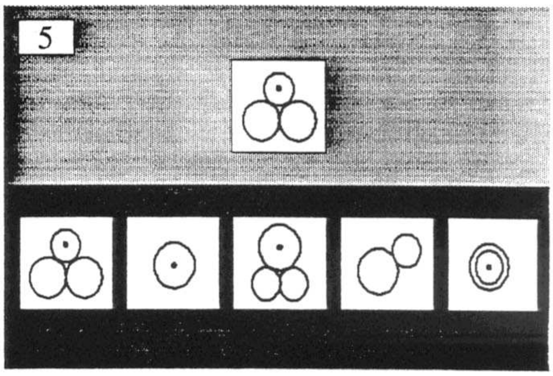


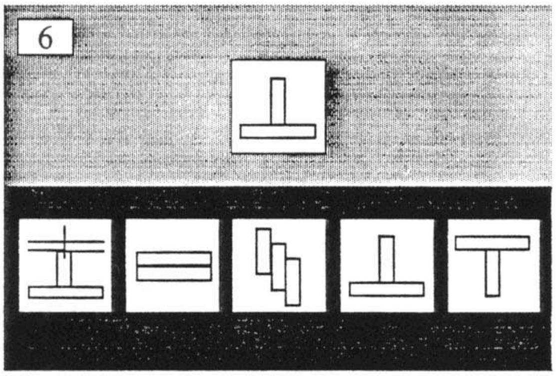


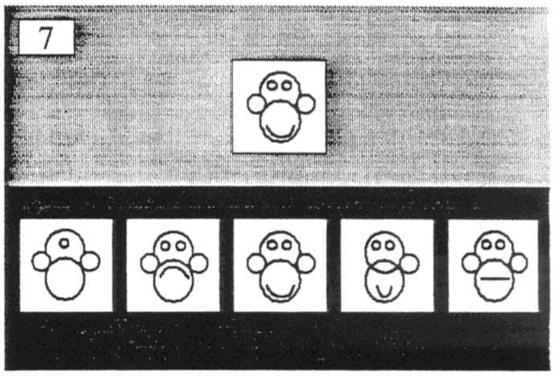


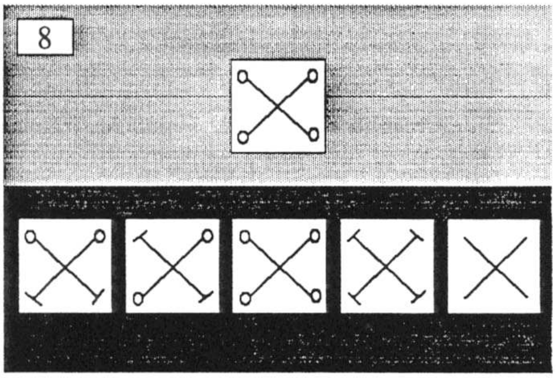
BaseCog Dokumentation: Gleiche Bilder 14.6.1990 – 2 –

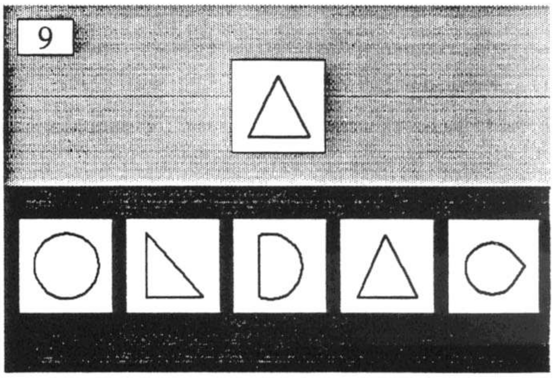


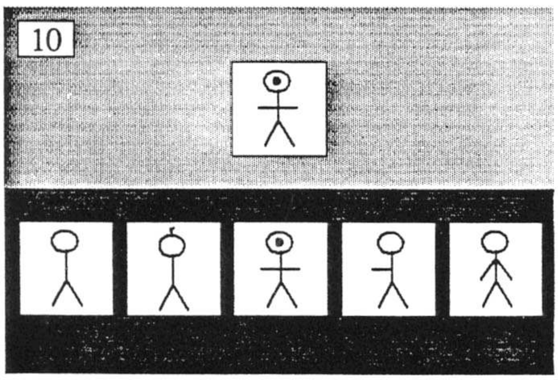


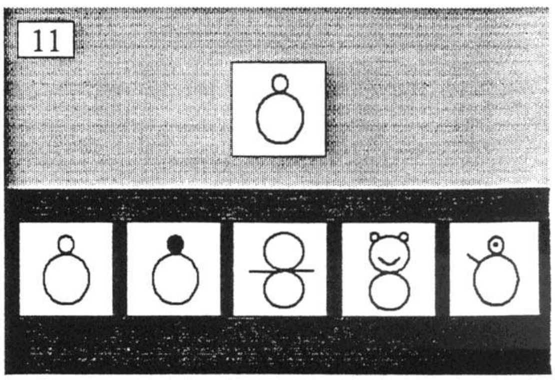




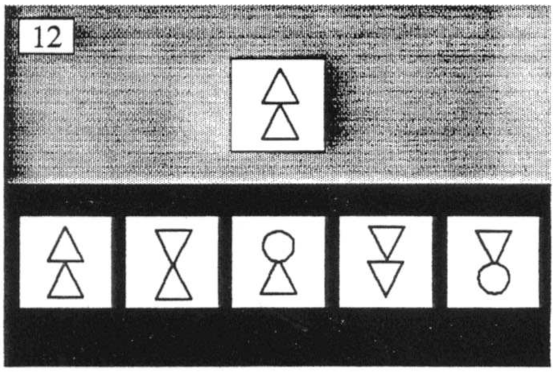


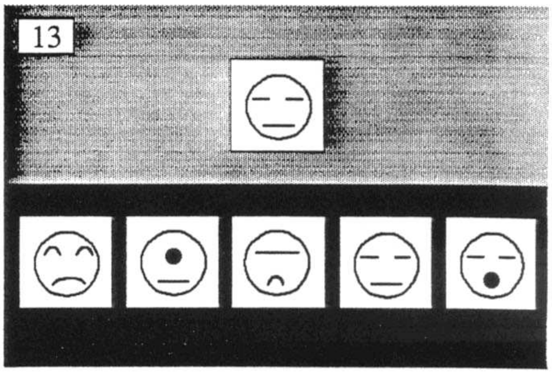


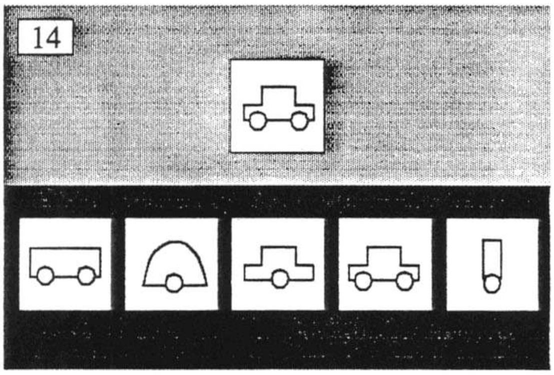


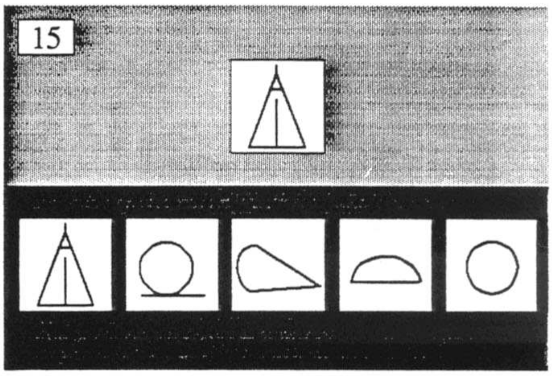


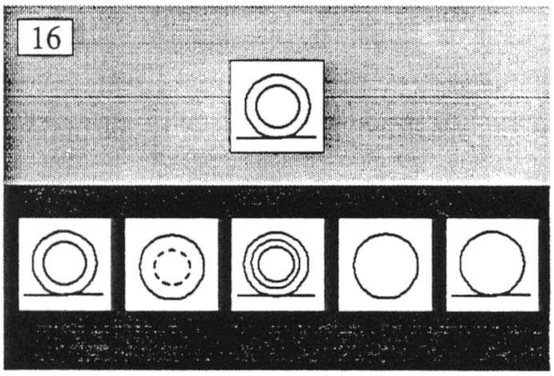
BaseCog Dokumentation: Gleiche Bilder 14.6.1990 – 3 –

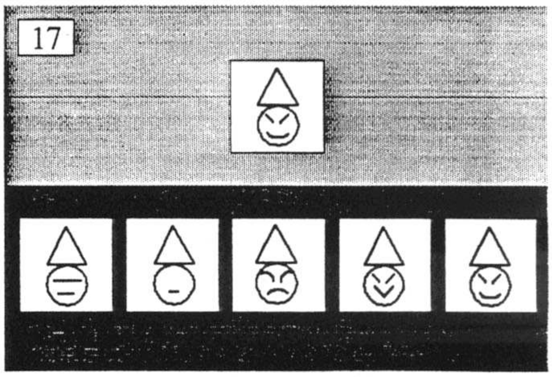


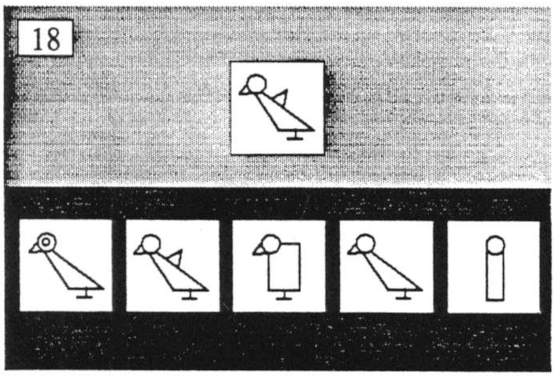


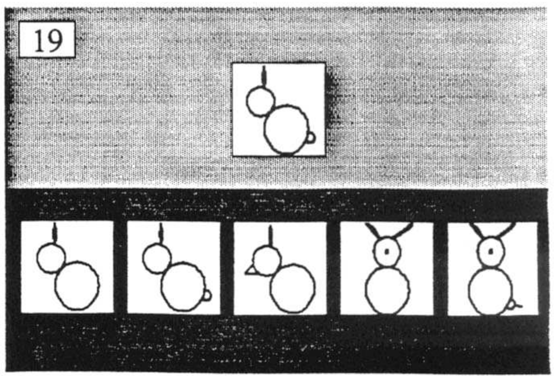




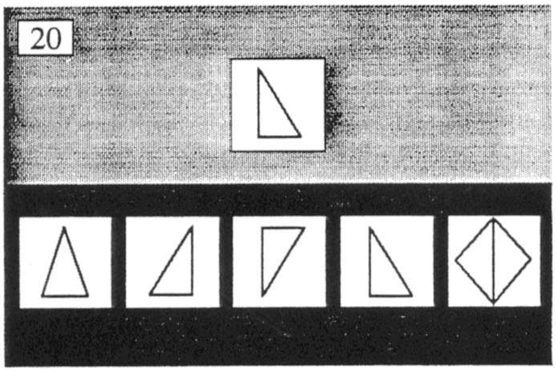


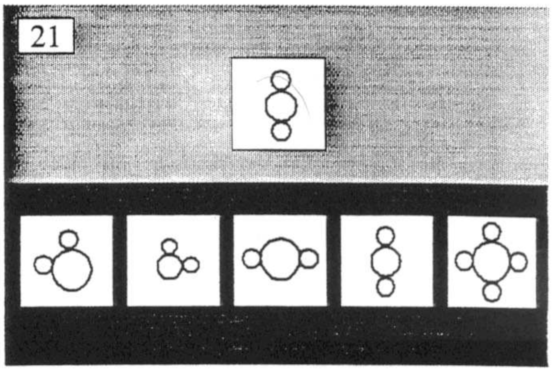


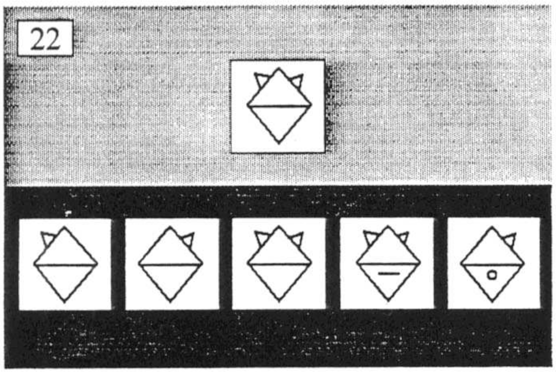


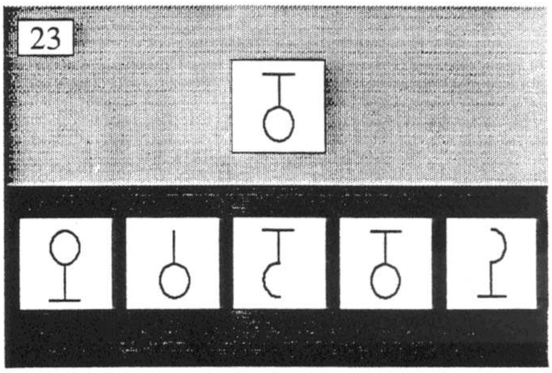


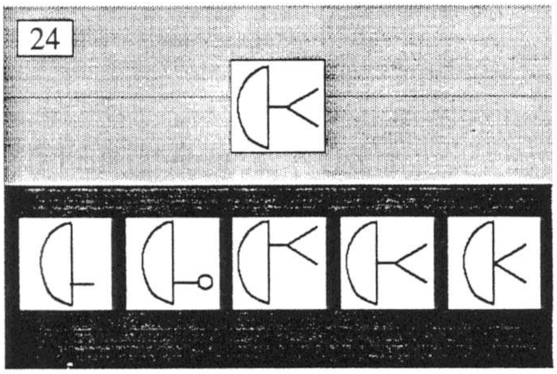
BaseCog Dokumentation: Gleiche Bilder 14.6.1990 – 4 –

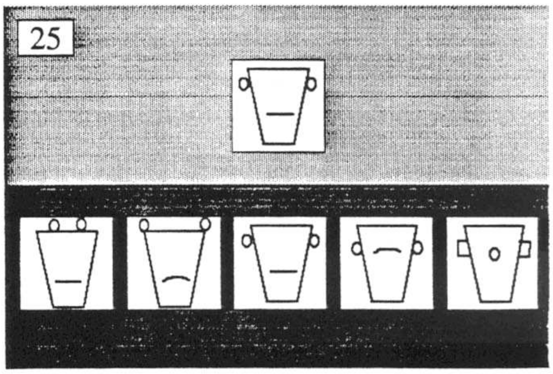


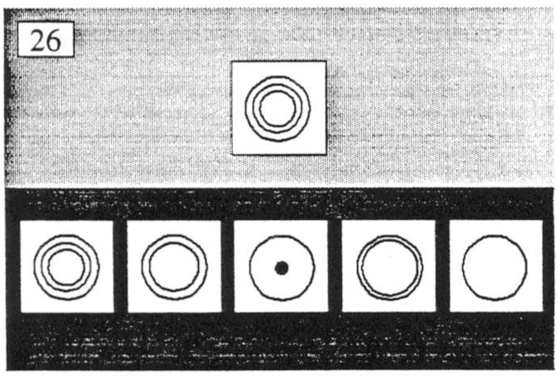


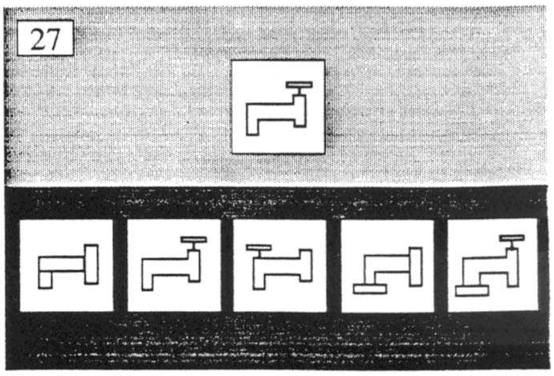




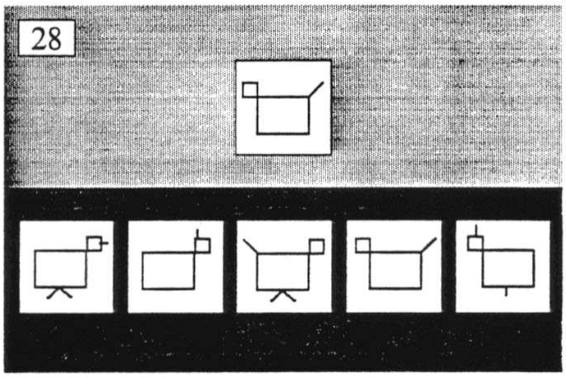


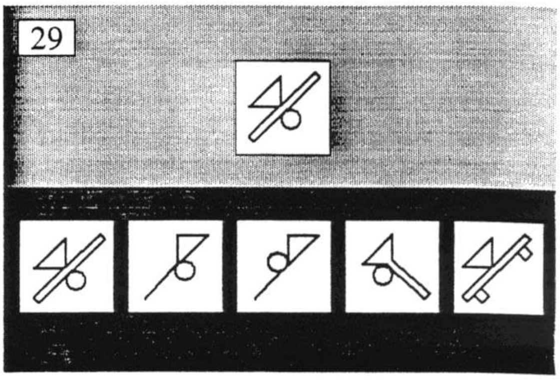


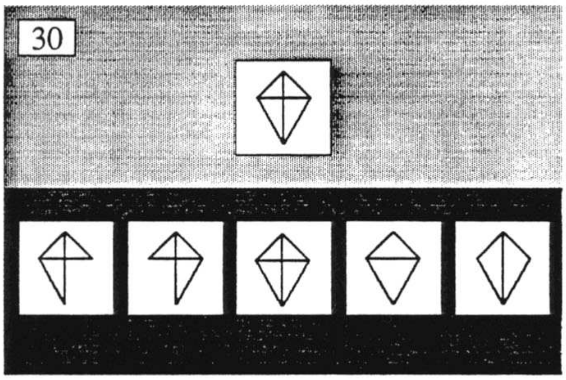


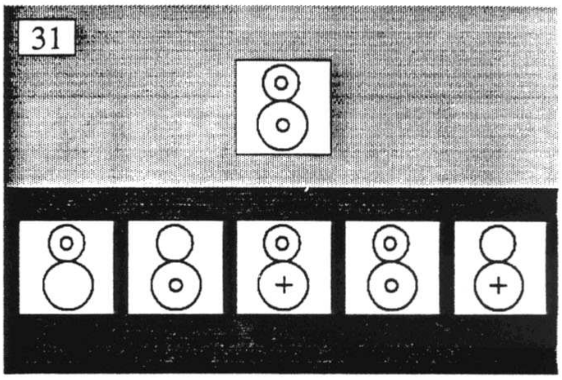


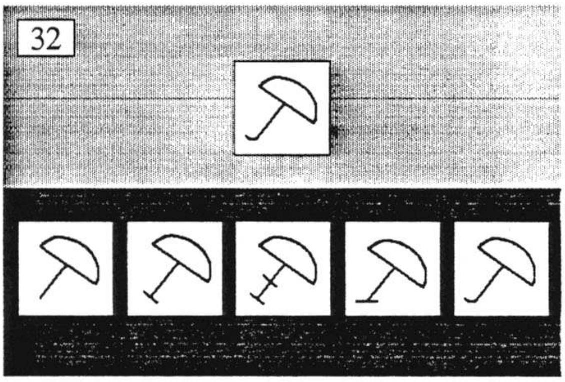
BaseCog Dokumentation: Gleiche Bilder 14.6.1990 – 5 –











Ende der Aufgabe

Memory: Memory for Text

ERKLÄRUNG

Bei der nächsten Aufgabe lese ich Ihnen eine kurze Geschichte vor. Danach werde ich Ihnen Fragen zum Inhalt der Geschichte stellen. Hören Sie also bitte gut zu, wenn ich Ihnen die Geschichte vorlese.

Gestern ging Peter, der sieben Jahre alt ist, zum Bleisee, um zu angeln. Er nahm seinen Hund Prinz mit. Der See war nach den starken Regengüssen der letzten vier Tage über die Ufer getreten.

Peter rutschte auf dem glitschigen Lehmboden aus und fiel in das tiefe Wasser. Er wäre ertrunken, wenn ihm nicht sein Hund nachgesprungen wäre und ihm geholfen hätte, das Ufer wieder zu erreichen.

Jetzt stelle ich Ihnen einige Fragen zum Inhalt der Geschichte.

Worum ging es im wesentlichen in dieser Geschichte ?

Wie hieß der Junge ?

Wie alt war der Junge ?

Wie hieß der Hund ?

Wie hieß der See ?

Warum fiel der Junge ins Wasser ?

Ende der Aufgabe

Memory: Paired Associates

BaseCog Dokumentation: Paarverbindungslernen 14.6.1990 – 1 –

ERKLÄRUNG
Wir zeigen Ihnen jetzt eine Reihe von Wortpaaren. Bitte versuchen Sie, sich diese Wortpaare zu merken. Anschließend erscheint auf dem Bildschirm jeweils das erste Wort eines Wortpaares und Sie sollen sich an das zweite Wort erinnern.

Wenn das Wortpaar also

Baum - Lampe

lautet, dann wird Ihnen später

Baum - ?

gezeigt und Sie sollen sich an "LAMPE" erinnern.

Wenn Sie jetzt START berühren, sehen Sie hintereinander eine Reihe von Wortpaaren, die Sie sich merken sollen.

START ⇨

ACHTUNG!

Familie - Zeitung

Familie - ?

Die gleichenWortpaare werden Ihnen nun noch einmal gezeigt. Bitte versuchen Sie, sich möglichst viele Wortpaare zu merken.

START ⇨

Familie - Zeitung

Knowledge: Spot-A-Word

BaseCog Dokumentation: Wörter finden 14.6.1990 – 1 –

ERKLÄRUNG

Bei den nächsten Aufgaben zeigen wir Ihnen jeweils fünf Wörter. Nur eines dieser Wörter ergibt einen Sinn. Dieses Wort sollen Sie finden.

Beispiel 1

- Bamu
- Paum
- Buma
- Baum
- Baun

Beispiel 2

- Purk
- Brug
- Lurk
- Krub
- Burg

Beispiel 3

- Maschine
- Mischane
- Schamire
- Meschina
- Manischa

Bitte beantworten Sie die folgenden Aufgaben so SCHNELL WIE MÖGLICH ! Wenn Sie jetzt den Pfeil mit START berühren, beginnt die Aufgabenreihe.

START ⇨

1

- siziol
- salzahl
- sozihl
- sziam
- sozial

2

- Lamone
- Talane
- Matrone
- Tarone
- Malonte

3

- kurinesisch
- kulinarisch
- kumenisch
- kulissarisch
- kannastrisch

BaseCog Dokumentation: Wörter finden 14.6.1990 – 2 –

4

- Tinxur
- Kukutur
- Fraktan
- Tinktur
- Rimsuhr

5

- beralen
- gerältet
- anälteren
- untären
- verbrämen

6

- Rosto
- Torso
- Soro
- Torgos
- Tosor

7

- Redor
- Radium
- Terion
- Dramin
- Orakium

8

- Nale
- Sahe
- Nase
- Nesa
- Sehna

9

- Rilkiase
- Kilister
- Riliker
- Klistier
- Linkure

10

- kentern
- knerte
- kanzen
- kretern
- trekern

11

- Sorekin
- Sarowin
- Rosakin
- Narosin
- Kerosin

BaseCog Dokumentation: Wörter finden 14.6.1990 – 3 –

12
- Ralke
- Korre
- Ruckse
- Recke
- Ulte

13
- Kapaun
- Paukan
- Naupack
- Aupeck
- Ankepran

14
- Dissonanz
- Diskrisanz
- Distranz
- Dinotanz
- Sidonenz

15
- Kirse
- Sirke
- Krise
- Krospe
- Serise

16
- Tume
- Umat
- Maut
- Taum
- Muta

17
- Tuhl
- Lar
- Lest
- Dall
- Lid

18
- Adept
- Padet
- Edapt
- Epatt
- Taped

19
- Spirine
- Saprin
- Parsin
- Purin
- Asprint

Knowledge: Vocabulary

BaseCog Dokumentation: HAWIE Wortschatz 14.6.1990 – 1 –

ERKLÄRUNG
Bei der nächsten Aufgabe geht es darum, wie viele Wörter Sie kennen. Auf dem Bildschirm sehen Sie immer ein Wort. Ihre Aufgabe besteht darin, seine Bedeutung zu erklären.

Beispiel 1

Was ist ein Apfel?

Beispiel 2

Was ist ein Esel?

Jetzt kommen weitere Wörter, die Sie erklären sollen. Mit dem Berühren von START können Sie beginnen.

START

1 Was heißt schleunigst ?

2 Was ist ein Gewebe ?

3 Was ist ein Opal ?

4 Was ist ein Pelz ?

5 Was ist ein Mandant ?

6 Was ist eine Nadel?

7 Was ist eine Deformation ?

8 Was ist Prestige ?

9 Was ist ein Scheck ?

10 Was heißt kriechen ?

11 Was ist ein Kurs ?

12 Was ist eine Oase ?

13 Was ist eine Andeutung ?

14 Was ist eine Sicht ?

15 Was ist Resistenz ?

16 Was ist ein Parlament ?

Ende der Aufgabe

B. 1. 2 Personality and Self-Related Functioning

Neuroticism

Original items	Translation used in BASE
When I'm under a great deal of stress, sometimes I feel like I'm going to pieces.	Wenn ich unter starkem Streß stehe, fühle ich mich manchmal, als ob ich zusammen-bräche.
I rarely feel fearful or anxious.	Ich empfinde oft Furcht oder Angst. *
I often feel tense and jittery.	Ich fühle mich oft angespannt und nervös.
I often get angry at the way people treat me.	Ich ärgere mich oft darüber, wie mich andere Leute behandeln.
I often feel helpless and want someone else to solve my problems.	Ich fühle mich oft hilflos und wünsche mir jemanden, der meine Probleme löst.
Sometimes I feel completely worthless.	Manchmal fühle ich mich völlig wertlos.

Note. * Double-negatives were avoided in the translation.

Extraversion

Original items	Translation used in BASE
I like to have a lot of people around me.	Ich habe gerne viele Leute um mich herum.
I really enjoy talking to people.	Ich unterhalte mich wirklich gerne mit anderen Menschen.
I like to be where the action is.	Ich stehe gerne im Mittelpunkt.
I laugh easily.	Ich bin leicht zum Lachen zu bringen.
I am a cheerful, high-spirited person.	Ich bin ein fröhlicher, gut gelaunter Mensch.
I am a very active person.	Ich bin ein sehr aktiver Mensch.

Internal Control Beliefs

BASE items / variables	English translation
Das, was an Gutem in meinem Leben passiert, kann **ich selbst** bestimmen.	It 's up to me to arrange for all the good things in my life.
Das Gute und Schöne in meinem Leben kann **ich selbst** beeinflussen.	I can make sure that good things come my way.
Wenn ich bekomme, was ich will, so ist das meistens, weil **ich selbst** viel dafür getan habe.	When I get what I want, it is usually because I have worked hard for it.

External Control Beliefs

BASE items / variables	English translation
Im allgemeinen sorgen **andere Leute** dafür, daß in meinem Leben nichts schief geht.	Other people generally make sure that nothing goes wrong in my life.
Ich bin auf **andere** angewiesen, um Unannehmlichkeiten zu vermeiden.	I depend on others to ensure that there are no problems in my life.
Zum größten Teil sorgen **andere Leute** dafür, daß in meinem Leben alles gut geht.	Other people generally arrange for good things to happen in my life.
Die angenehmen Dinge in meinem Leben hängen von **anderen Leuten** ab.	The good things in my life are determined by other people.

Personal Life Investment

BASE items / variables	English translation
Wie sehr denken Sie gegenwärtig daran oder tun etwas im Zusammenhang damit?	How much do you think about it or do something for it?
Also wie ist es mit ...	So, how about your...
Ihrer Gesundheit	Your health
Ihrer geistigen Leistungsfähigkeit (z.B. Ihr Gedächtnis)	Your cognitive fitness
Ihren Hobbys und anderen Interessen	Your hobbies and interests
der Beziehung zu Ihren Freunden und Bekannten	Your relations with friends and acquaintances
Ihrer Sexualität	Your sexuality
dem Wohlergehen Ihrer Angehörigen	The well-being of your family
Ihrer beruflichen oder einer anderen vergleichbaren Tätigkeit	Your professional or a comparable activity
Ihrer Unabhängigkeit	Your independence
dem Nachdenken über Ihr Leben	Thinking about your life
Ihrem Sterben oder Ihrem Tod	Your death and dying

B. 1. 3 Social Integration

Emotional Loneliness

Original items	Translation used in BASE
I do not feel alone.	Ich fühle mich allein (r). *
I lack companionship.	Ich habe wenig Gesellschaft (r).
I feel isolated.	Ich fühle mich isoliert. (r)
I feel left out.	Ich fühle mich ausgeschlossen. (r)

Note. * Double-negatives were avoided in the translation. (r) recoded.

Social Loneliness

Original items	Translation used in BASE
There are people I feel close to.	Es gibt Menschen, die mir nahe stehen.
There are people I can turn to.	Es gibt Personen, an die ich mich vertrauensvoll wenden kann.
I feel part of a group of friends.	Ich fühle mich einem Bekanntenkreis zugehörig.
There are people I can talk to.	Es gibt Menschen, mit denen ich offen sprechen kann.

Number of Close Others

Original item	Translation used in BASE
Is there any one person or persons that you feel so close to that it's hard to imagine life without them? (first circle)	In den ersten Kreis, der Ihnen also am nächsten liegt, sollen alle die Leute aufgenommen werden, denen Sie sich so eng verbunden fühlen, daß Sie sich ein Leben ohne sie nur schwer vorstellen können.

B. 2 Cross-Disciplinary Measures

Well-Being: Non-Agitation

Original items	**Translation used in BASE**
Little things bother me more this year.	Dieses Jahr rege ich mich über Kleinigkeiten auf.
I sometimes worry so much that I can't sleep.	Ich mache mir oft solche Sorgen, daß ich nicht einschlafen kann.
I am afraid of a lot of things.	Ich habe vor vielen Dingen Angst.
I get mad more than I used to.	Ich werde häufiger wütend als früher.
I take things hard.	Ich nehme die Dinge schwer.
I get upset easily.	Ich rege mich leicht auf.

Well-Being: Aging Satisfaction

Original items	**Translation used in BASE**
Things keep getting worse as I get older.	Je älter ich werde, desto schlimmer wird alles.
I have as much pep as I had last year.	Ich habe noch genauso viel Schwung wie letztes Jahr. (r)
As you get older, you are less useful.	Je älter ich werde, desto weniger nützlich bin ich.
As I get older, things are better/ worse than I thought they would be.	Mit zunehmendem Alter ist mein Leben besser, als ich erwartet habe. (r)
I am as happy now as when I was younger.	Ich bin jetzt genauso glücklich, wie ich es in jungen Jahren war. (r)

Note. (r) recoded.

Well-Being: Life Satisfaction

Original items	**Translation used in BASE**
I sometimes feel that life isn't worth living.	Manchmal glaube ich, daß das Leben nicht lebenswert ist.
Life is hard for me most of the time.	Das Leben ist die meiste Zeit hart für mich.
I have a lot to be sad about.	Ich bin über vieles traurig.
How satisfied are you with your life today?	Wie zufrieden sind Sie zur Zeit mit ihrem Leben? (r)

Note. (r) recoded.

B. 3 Reliabilities Over Time for Measures of Personality, Self-Related Functioning, and Social Integration

Table B.1 reports reliability coefficients for measures of personality and self-related functioning, and social integration. Reliabilities for the cognitive measures have widely been documented (e.g., Lindenberger et al., 1993), and the circle task to assess the number of close confidants represents a single item and its characteristics have also been documented (e.g., Antonucci, 1986). In line with other BASE reports (e.g., Smith & Baltes, 1997; 1999), all analyses of the present dissertation that were carried out with the measures used in Table C.1 (with the exception of goal investment) were based on item parcels rather than raw mean scores of the items. The rationale underlying parceling has been described in Section 3.2.1.1 and details about the parceling procedure can be obtained from P. B. Baltes and Mayer (1999). Test-retest correlations for the measures used in the dissertation can be obtained from Table 6 in Section 3.2.1.1.

Table B. 1
Reliabilities for Measures of Personality, Self-Related Functioning, and Social Integration

	Cronbach's Alpha		
Measure	Baseline	2nd Wave	3rd Wave
Neuroticism	.77	.74	.74
Extraversion	.62	.63	.68
Internal Control	.54	.57	.40
External Control	.75	.73	.78
Goal Investment	.61	.65	.72
Social Loneliness	.75	.68	.70
Emotional Loneliness	.75	.75	.71

Note. *N* = 130.

APPENDIX C RESULTS

C. 1 Identification and Description of Heterogeneity and Its Correlates: Subgroup Differences in Psychological Profiles and External Correlates at Baseline

C. 1. 1 Results of Cluster Analysis: Comparison to Previous Systemic-Wholistic Work From the Berlin Aging Study

A central criterion to determine the adequacy of the three subgroups identified in the 6-year longitudinal BASE sample was to link these subgroups to the desirability profile status established for subgroups from the total cross-sectional BASE sample (Smith & Baltes, 1997). The following section complements analyses reported in Section 4.1.2, which linked the three subgroups to the pooled desirable and less desirable profile subgroups from the total BASE sample. Below, it is asked to what extent persons who were grouped together by cluster analysis in the total BASE sample have also been grouped together by the cluster analysis when the much smaller measurement space of the longitudinal sample was used. Both cluster analyses were carried out at baseline assessment.

Panel A of Figure C.1 shows the membership distribution for participants who were previously classified as showing desirable profiles (Smith & Baltes, 1997). It has to be acknowledged that only those participants are considered who took part three times in BASE. It can be seen, for example, that most participants from Subgroup 1 (tentative label used by Smith and Baltes, 1997: Cognitively very fit and vitally involved) and Subgroup 2 (socially oriented and involved) that were identified by Smith and Baltes (1997) in the total BASE sample were classified into the Overall-Positive Profile Subgroup from the longitudinal BASE sample. Longitudinal participants in the relatively large Subgroup 3 (cognitively fit, well balanced) were more or less split up on the three newly extracted subgroups with the largest percentage being classified into the Overall-Positive Profile Subgroup. All members of the previously identified Subgroup 4 (cognitively fit, but reserved loner) were re-grouped into the Disparate-Profile Subgroup.

Panel B of Figure C.1 shows the membership distribution for participants from the previously classified less functional/desirable profiles (Smith & Baltes, 1997). Only 24 out of 270 participants in these subgroups were available for repeated assessment over six years. As a consequence, the frequency of participants in these five subgroups who remained in BASE over time was relatively small. The seven participants who remained in BASE from Subgroup 5 (fearful, lonely, but supported) were more or less equally split up on the newly extracted Average-Profile Subgroup and Disparate-Profile Subgroup. Of the 11 longitudinal members from Subgroup 6 (anxious, lonely, holding on to control), the majority was classified into the Disparate-Profile Subgroup, whereas all longitudinal participants from Subgroup 7 (n = 2; impaired, dependent, but well-balanced) and Subgroup 8 (n = 4; cognitively impaired, disengaged, but content) were categorized into the newly extracted Average-Profile Subgroup. All participants from the least well-functioning subgroup in the total BASE sample (Subgroup 9; impaired, withdrawn, in despair) completely dropped out over time.

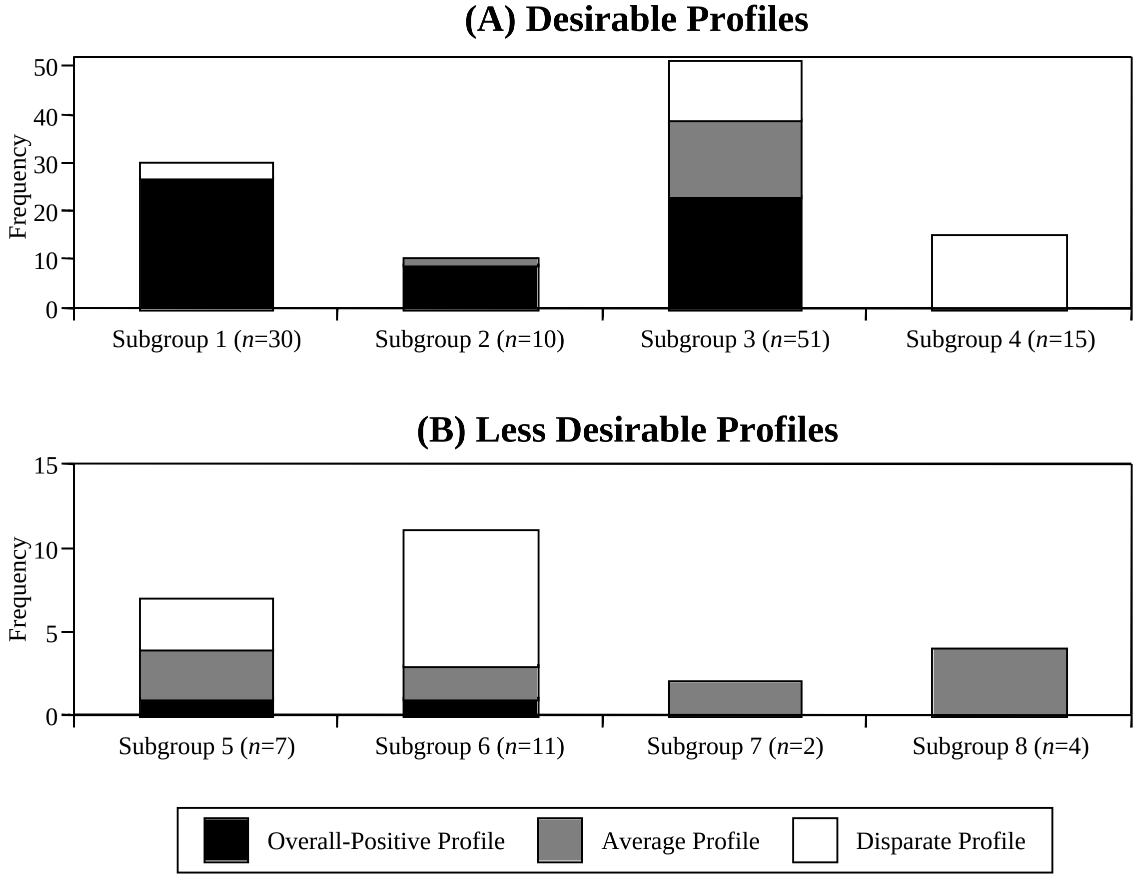


Figure C. 1 Description of the three subgroups identified in the 6-year longitudinal BASE sample in terms of the desirability profile status established for the nine subgroups from the total BASE sample.

Note. The large majority of the 6-year longitudinal sample was among the desirable profile subgroups identified in the total cross-sectional BASE sample. Note that Smith and Baltes (1997) identified four subgroups with desirable profiles and five subgroups with less desirable profiles. No participant from Subgroup 9 identified in the total BASE sample remained in the study over six years.

C. 1. 2 Subgroups Extracted From the 6-Year Longitudinal BASE Sample: Profile Characteristics on Psychological Measures at Baseline Assessment

The following section reports additional analyses of differences between the three subgroups identified at baseline assessment of the 6-year longitudinal sample. Specifically, differences in within-subgroup homogeneity and the contribution of single profile-defining measures to overall group separation were considered.

Over and above mean level differences, the three subgroups also differed from one another in within-subgroup homogeneity on four out of the 11 clustering measures. Table 12 (see Section 4.1.3) provides variability information for the three subgroups across the measures. Levene's test for homogeneity of variances indicated that the Average-Profile Subgroup showed the largest amount of variability on knowledge, F (2,127) = 4.56, $p < .05$; the Disparate-Profile Subgroup was most heterogeneous on emotional loneliness, F (2,127) = 3.71, $p < .05$, and social loneliness, F (2,127) = 9.43, $p < .001$; and on number of close others it was the Overall-Positive Profile Subgroup that showed relatively little intra-group homogeneity, F (2,127) = 4.55, $p < .05$. The overall patterning suggests that the subgroups did not differ very much from each other in their intra-group homogeneity, and there was not a single subgroup that was consistently more heterogeneous or less heterogeneous than other subgroups.[1]

In addition, discriminant analysis was not only used to provide convergent evidence for reliability of subgroup assignment (albeit from a multi-linear perspective; see Section 4.1.1), but also to confirm the most distinguishing profile-defining measures (i.e., which variables contribute most to subgroup separation). One of the two canonical functions that discriminate the subgroups from one another in discriminant analysis largely represented the social integration domain (loadings for emotional loneliness and social loneliness amounted to .50) and the other one mainly represented the cognitive domain, particularly crystallized aspects (loading for knowledge amounted to .77). In accord with the ANOVA results, correlations found between the profile-defining and the canonical functions at baseline assessment showed that external control was the only profile-defining variable without discriminatory power.[2] From the discriminant analysis follows that measures of social integration such as loneliness and measures of cognitive functioning such as crystallized intelligence contributed most to separating the three subgroups from one another, at least from this analysis perspective.

[1] Heterogeneity of variance across the groups may violate the assumptions underlying a proper application of statistical tests used in the present dissertation (e.g., ANOVA). Analyses aimed at checking the assumptions of statistical tests used in this study are reported in the next Section.
[2] However, external control beliefs had statistically significant discriminatory power at both follow-up occasions.

C. 1. 3 Subgroups Extracted From the 6-Year Longitudinal BASE Sample: Profile Characteristics on Cross-Disciplinary Measures at Baseline Assessment

Table C.1 presents cross-disciplinary characteristics of the there subgroups. Demographic information as well as mean level and variability for measures of physical functioning and life-history factors are given. Three sets of findings of particular interest.

First, although not statistically significant, there was a numerical tendency for the Disparate-Profile Subgroup to contain more women and more persons living alone. For example, only 43% of the Average-Profile Subgroup were women; for the Disparate-Profile Subgroup, the ratio was 66%.

Second, contrast analyses revealed that the Overall-Positive Profile Subgroup was significantly better-off than the Disparate-Profile Subgroup on (a) three out of the four physical functioning measures (vision, multimorbidity, and ADL-IADL), and (b) one out of three life-history measures (education). Similarly, participants from the Overall-Positive Profile Subgroup were significantly better off than the Average-Profile Subgroup on six out of seven measures (hearing, vision, multimorbidity, ADL/IADL, education, and prestige). There were comparably fewer statistically significant differences between the Average-Profile Subgroup and the Disparate-Profile Subgroup. The only difference found was that the Disparate-Profile Subgroup had a higher occupational prestige score.

Third, data summarized in Table C.1 also shows that the subgroups differed in within-group homogeneity on some of the measures. Levene's test for homogeneity of variances indicated that participants in the Overall-Positive Profile Subgroup not only had fewer limitations in (instrumental) activities of daily living, but also showed very little variation on this measure, $F(2,127) = 7.70$, $p < .001$. In a similar vein, the Average-Profile Subgroup was socio-economically disadvantaged in terms of education and professional prestige, and members of this group were much more homogeneous on these constructs as compared to both other subgroups, $F(2,127) = 3.57$, $p < .05$; and $F(2,127) = 5.93$, $p < .001$, respectively. The Disparate-Profile Subgroup also showed more variability on multimorbidity, $F(2,127) = 5.62$, $p < .01$.[3]

[3] The same results were obtained when Bartlett's Box test for homogeneity of variances was used. Here, subgroup differences in variability were also found for income.

Table C. 1
Cross-Disciplinary Characteristics of the Three Subgroups Identified at Baseline Assessment of the 6-Year Longitudinal BASE Sample

Measure	Subgroups					
	Overall-Positive Profile (*n* = 61)		Average Profile (*n* = 28)		Disparate Profile (*n* = 41)	
Demographics						
% Women	53		43		66	
% Living Alone	57		57		71	
	M	*SD*	*M*	SD	*M*	*SD*
Age	76.6_{a}	$\mathit{4.5}_{a}$	78.4_{a}	$\mathit{5.1}_{a}$	80.7_{b}	$\mathit{7.6}_{b}$
Physical Functioning						
Hearing	58.4_{a}	$\mathit{9.8}_{a}$	50.8_{b}	$\mathit{10.6}_{a}$	$55.1_{a,b}$	$\mathit{8.8}_{a}$
Vision	58.7_{a}	$\mathit{8.4}_{a}$	53.0_{b}	$\mathit{8.6}_{a}$	54.3_{b}	$\mathit{9.6}_{a}$
Multimorbidity *	57.8_{a}	$\mathit{6.5}_{a}$	53.6_{b}	$\mathit{8.5}_{a}$	49.4_{b}	$\mathit{11.3}_{b}$
ADL/IADL	56.9_{a}	$\mathit{1.8}_{a}$	54.9_{b}	$\mathit{6.1}_{b}$	54.4_{b}	$\mathit{5.4}_{b}$
Life History						
Education	55.8_{a}	$\mathit{10.7}_{a}$	46.5_{b}	$\mathit{7.0}_{b}$	50.5_{b}	$\mathit{8.5}_{a}$
Occupational Prestige	53.6_{a}	$\mathit{11.0}_{a}$	45.2_{b}	$\mathit{5.5}_{b}$	54.3_{a}	$\mathit{10.3}_{a}$
Income	51.3_{a}	$\mathit{9.1}_{a}$	47.6_{a}	$\mathit{6.8}_{a}$	53.7_{a}	$\mathit{13.9}_{a}$

Note. * Scores on this dimension were reverse-coded to calculate the desirability (functional status) score. Means (*M*) and standard deviations (*SD*), respectively, in the same row that do not share subscripts differ at the $p < .05$ level or below.

C. 1. 4 Checking Assumptions of Statistical Tests

The following paragraphs summarize analyses aimed at investigating whether or not the statistical assumptions associated with an adequate application of the statistical tests used in the present study were fulfilled. Many of these tests share the same assumptions so that the paragraphs are centered around the assumptions rather than around the tests: (1) multivariate normality, (2) homogeneity of variance-covariance matrices, (3) homogeneity of covariances between repeated assessments, (4) independence of observations, (5) multicollinearity and singularity, and (6) linearity and homoscedasticity.

Multivariate normality is a central assumption in applications of many statistical tests including (repeated measures) ANOVAs and structural equation modeling. Univariate normality can be regarded a necessary, but insufficient condition for multivariate normality. In the present investigation, the assumption of univariate normality was approximately met by all measures at each occasion: If skewness was present, it was at maximum 1.46 (number of close others at the last occasion), which indicates that skewness was not extreme. In practice, Tabachnick and Fidell (1996) recommend a rule of thumb: "*Unless there are fewer cases than dependent variables (DVs) in the smallest group and highly unequal* n*, deviation from normality of sampling distributions is not expected* (p. 444)." Although the requirement of equal group sizes is only partly met, Mardia (1971) demonstrated that robustness of a (repeated measures) ANOVA was assured even in case of unequal numbers within each subgroup as long as the smallest cell is sufficiently large ($N > 20$). Similarly, Mardia (1971) showed that a (repeated measures) ANOVA is robust to modest violations of assumptions of multivariate normality if the violation is created by some degree of skewness rather than by outliers.[4] On this basis, multivariate normality of the sampling distribution of means can be expected in the present study.

Another assumption relates to the *homogeneity of variance-covariance matrices*. Levene's test for homogeneity of variances and Box's *M* multivariate test for homogeneity of dispersion matrices were used to test this assumptions on measures of psychological functioning as well as the cross-disciplinary factors. Single variables were found to violate the homogeneity assumption. For example, Levene's test indicated that the Overall-Positive Profile Subgroup was more heterogeneous on number of close others, (2,127) = 4.55, $p < .05$. Similarly, Box's *M* test was found statistically significant for the profile-defining psychological measures at baseline assessment, Box's M= 202.71, F (132, 23082) = 1.32, $p < .05$. However, several notes of caution appear necessary when interpreting these violations. First, the outcome of Box's *M* test can be safely disregarded (unless it is below $p < .001$) because this test is notoriously sensitive (Tabachnick & Fidell, 1996). Second, it was also found that (a) differences in variances and covariances were not extreme (i.e., much less than the critical 10:1 ratio), (b) none of the subgroups was consistently more heterogeneous than others, and (c) none of the variables showed heterogeneous variance-covariance matrices consistently over time.

Third, Tabachnick and Fidell (1996) concluded from the existing literature on robustness that the alpha level is conservative if larger samples produce larger variances and covariances. The significance is too liberal only if smaller sample sizes produce larger (co-)variances. In the context of the present study, the group with the smallest sample size was the Average-Profile Subgroup ($n = 28$). This group of persons was not found to show differential

[4] Similarly, outliers have a relatively larger impact on the size of the regression coefficient. In analogy to the Smith and Baltes study (1997), two multivariate outliers were not considered in the present analyses.

change over time, and it was only on knowledge at baseline assessment where the Average-Profile Subgroup was found to produce the largest variances (see Section 4.1.3 for Table 12; Overall-Positive Profile: *SD* = 5.9; Average-Profile: *SD* = 8.3; Disparate-Profile: *SD* = 5.0). Finally, it has to be acknowledged that ADL/IADL was not considered in the analyses because this measure was found to show ceiling effects in the Overall-Positive Profile Subgroup (see Table C.1). In summary, there were some indications that homogeneity of variances and covariances were violated in the present study, but (repeated measures) ANOVAs are reasonably robust in this regard. In addition to Wilk's λ, Pillai's criterion was also examined because this coefficient has been demonstrated to be a more robust index of multivariate significance (Olson, 1979). In none of the analyses carried out, Pillai's criterion and Wilk's λ differed from one another in deciding about rejecting the null hypothesis.

One further central assumption of a repeated measures ANOVA concerns the *homogeneity of covariances between repeated assessments*. The assumption states that all pair-wise correlations among the levels of the within-subject variable(s) have to be approximately equivalent. In the context of the present study, meeting the assumption may be particularly problematic because some pairs of levels (i.e., 2-year interval between the second wave and the third wave) are closer in time, while other pairs are more distant in time (i.e., 4-year interval between baseline assessment and the second wave). This violation results in an inflated Type I error rate (i.e., tests are not conservative enough). Among the approaches used to compensate this potential flaw are the following options: (a) use a reasonably conservative F test with $(1, n - 1)$ degrees of freedom, (b) adjust the degrees of freedom by a factor that reflects the amount of the violation of homogeneous covariances. Statistics computer packages such as SPSS and SAS routinely provide adjusted means and the Huynh-Feldt and Greenhouse-Geisser corrections for failure of the assumption. In all of the analyses carried out, these corrections did not alter the findings reported. Reporting the significance of the multivariate (in addition to the univariate) effect also circumvents the problem because (transformed) multivariate dependent variables (DVs) replace the within-subjects independent variable (IV) and the assumption of homogeneity of variances is no longer required (Tabachnick & Fidell, 1996).

Independence of observations means that the scores of the participants were not influenced by one another. This assumption can be regarded to be met because every session was administered in a one-to-one testing situation at the participant's home, and testing was carried out by trained research assistants.

Lack of *multicollinearity and singularity* also plays a major role in many statistical tests such as ANOVAs and multiple regression analyses, both logically and statistically. This assumption relates to the amount of intercorrelations among variables. In (repeated measures) ANOVAs, high intercorrelations among the DVs imply that one DV is a linear combination of other DVs and thus provides information that is redundant to the information that other DVs can provide. In regression analyses, multicollinearity refers to intercorrelations among the IVs, which may cause serious problems. The greater the multicollinearity, the more unstable are (partial) regression coefficients, and the larger are standard errors and confidence intervals, which results in decreased likelihood of statistically significant results. In the current work, intercorrelations among the measures used were at maximum moderately high (see Section 4.1.1 for Table 9) and were considerably lower than the threshold of $r = .80$ (Licht, 1995).

Specific assumptions of regression analyses are *normality, linearity,* and *homoscedasticity* between predicted DV scores and errors of prediction. More specifically, the difference between obtained and predicted DV scores (residual) is expected to be normally distributed over the predicted DV scores, that residuals have a straight-line relationship with predicted DV scores, and that error scores have equal variances at all values of the predictors (Licht, 1995; Tabachnick & Fidell, 1996). All sets of regression analyses were accompanied by graphically examining the assumptions, and none of these analyses appeared to reveal strong deviations from the above conditions. The literature suggests that moderate violations tend not to be very problematic (see Pedhazur, 1982) because failure of these assumptions does not invalidate an analysis so much as weakens it. For example, homoscedasticity reduces the statistical power to detect a significant effect when the effect does truly exist in nature.

In sum, results of a preliminary check of the statistical assumptions revealed that all of the above mentioned assumptions were at minimum roughly met. This provides the methodological basis for applying the statistical tests used in the present study.

C. 2 Differential Development Over Time and Its Underlying Mechanisms: Subgroup Differences in Change of Psychological Profiles Over Six Years and the Role of External Correlates

C. 2. 1 Stability and Change at the Level of the Subgroup-Defining Measures

Regression artifacts represent a general threat to research on change over time, and several different strategies have been suggested to adjust for regression artifacts (for overview, see Campbell & Kenny, 1999). One strategy was proposed by Nesselroade and colleagues (1980) who suggested to omit the first measurement occasion for change analyses and restrict change analyses to subsequent occasions. Implementing this approach in the context of the present study would have meant to use baseline data for classification purposes only, and data from the second wave and the third wave of BASE for change analyses. This recommendation is problematic insofar as information from one of three measurement occasions would have been lost, and the time interval to observe change would have been substantially reduced from six years to two years. Given this, it was decided to compare the results from three different methodological approaches for convergence. First, based on event history analysis methodologies rather than using difference scores, Schaie (1989, 1990) suggested to convert continuous change into discrete change, adjusting for whether observed change has exceeded possible measurement error. Second, analyses using Multilevel Growth Curve Modeling were carried out (for review, see Hertzog & Nesselroade, 2003; Ghisletta & Lindenberger, 2004). The next sections first provide a brief introduction into these methods and their advantages and then report results of these analyses.

C. 2. 1. 1 Schaie's Procedure to Adjust for Error of Measurement

As a first strategy to examine whether subgroup differences in change over time were more than a mere reflection of a regression artifacts, a procedure was used that has been proposed by Schaie (1989, 1990). Observed change is regarded to be reliable in case that the post-test score exceeds the standard error of measurement (SEM) confidence interval around the participant's pre-test score. The standard error of measurement is defined as follows: $\sigma_{11} (1 - r_{12}^2)^{1/2}$ where σ_{11} is the standard deviation of the pre-test score and r_{12} is the correlation between the pre- and the post-test score.

Because this statistic is based on the standard error of prediction rather than the conventional standard error of measurement (which would involve *not* to square reliability), confidence intervals for this procedure are larger (see Dudeck, 1979; Hsu, 1995). Given this, the procedure proposed can be regarded to be more conservative than conventional methods (Gulliksen, 1950).[5] The underlying rationale for this statistic is whether or not the post-test score could have been another estimate of the pre-test score, given a certain confidence interval around the pre-test score.

To use the approach, change scores between baseline assessment and the third wave of BASE were used (neglecting the second wave). The minimum difference between pre-test and post-test scores required to be regarded as reliable change was 3.7 *T* scores for cognition, 3.5 *T* scores for self and personality, 5.3 *T*-score units for social integration, and 2.6 *T* scores for the overall measure of functioning/desirability. Using these cut-off points, there were very few cases who showed an increase of functioning over time (cognition $n = 7$, self and personality $n = 12$, social integration $n = 12$, and overall functioning/desirability $n = 5$).

Table C.2 contrasts results from the repeated measures ANOVAs against those found when using the procedure proposed by Schaie (1989, 1990). Overall, convergent evidence for the patterning of subgroup differences in change over time was found. On the composite measure of cognitive functioning, a significantly larger percentage of participants from the Disparate-Profile Subgroup ($n = 15$; 37%) showed negative change trajectories as compared to both other subgroups (Overall-Positive Profile Subgroup: $n = 13$, 21%; Average-Profile Subgroup: $n = 2$, 7%; χ^2 (4, $N = 130$) = 13.37, $p < .01$). Also in accord with results from the repeated measures ANOVA's, it was found that more participants from the Overall-Positive Profile Subgroup were classified as declining on both composite measures of self-related functioning and personality ($n = 26$, 43%) and social integration ($n = 22$, 36%) as compared to members from both other subgroups (self and personality: Average-Profile Subgroup: $n = 7$, 25%, Disparate-Profile Subgroup: $n = 8$, 20%, χ^2 (4, $N = 130$) = 10.29, $p < .05$; social integration: Average-Profile Subgroup: $n = 4$, 14%, Disparate-Profile Subgroup: $n = 3$, 7%, χ^2 (4, $N = 130$) = 22.03, $p < .001$). Finally, on the overall functioning/desirability measure, more participants from the Overall-Positive Profile Subgroup ($n = 28$; 46%) showed negative change trajectories as compared to both other subgroups (Average-Profile Subgroup: $n = 7$, 25%; Disparate-Profile Subgroup: $n = 10$, 24%; χ^2 (4, $N = 130$) = 11.42, $p < .05$). Figure C.2 graphically illustrates that the percentage of participants within the subgroups who declined over six years differed across the three domains examined.

[5] Although this statistic is more conservative than others, Schaie (1989) argued that it still might be biased against finding stability over time because of an error rate of .16 in favor of accepting the presence of reliable change.

Table C. 2
Differences in Change Over Six Years Between The Three Subgroups Identified at Baseline Assessment of the 6-Year Longitudinal BASE Sample: Convergent Evidence from Repeated Measures ANOVAs (A) and Schaie Procedure (B)

	Overall Positive Profile		Average Profile		Disparate Profile	
Construct	A	B	A	B	A	B
Cognition	–	13 (21%)	–	2 (7%)		15 (37%)**
Self & Personality		26 (43%)*	–	7 (25%)	–	8 (20%)
Social Integration		22 (36%)***	–	4 (14%)	–	3 (7%)
Overall Desirability		28 (46%)*	–	7 (25%)	–	10 (24%)

Note. Numbers and percentages in column B refer to decline over time. ─ = Relative Stability, = Decline. $p < .05$; – ns.; * $p < .05$; ** $p < .01$; ** $p < .001$.

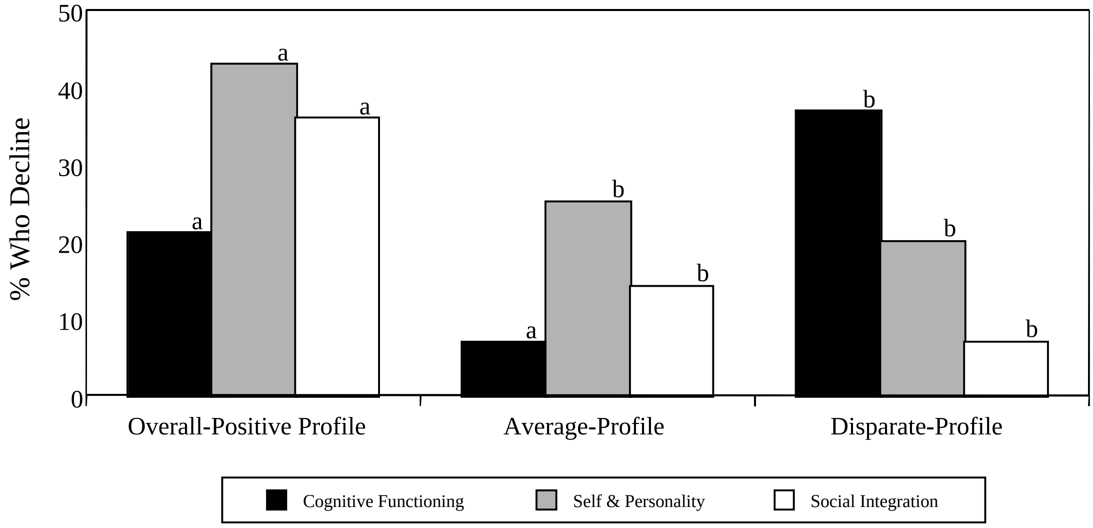


Figure C. 2 Subgroup decline differed across domains of psychological functioning: Results from Schaie's adjustment for error of measurement (1989).
Note. Within measures, subgroups with different superscripts differ at $p < .05$ or below.

Taken together, adjusting the change scores for unreliability based on a procedure recommended by Schaie (1989, 1990) revealed essentially the same pattern of results as those found in the repeated measures ANOVAs. Participants from the Overall-Positive Profile Subgroup were more likely to have shown negative change trajectories on non-cognitive variables as were members of both Average-Profile Subgroup and Disparate-Profile Subgroup. In contrast, members from the Disparate-Profile Subgroup were more likely to decline on cognitive measures as compared to members from both other subgroups.

C. 2. 1. 2 Multilevel Growth Curve Modeling

A further set of statistical techniques that may handle some of the methodological problems associated with longitudinal data is individual growth modeling (Bryk & Raudenbush, 1992; Nesselroade & Boker, 1994; Rogosa & Willett, 1985). Individual growth modeling is a type of multilevel models (MLM; also known as random effects models or hierarchical linear modeling). In a recent comparative study, Reynolds, Gatz, and Pedersen (2002) demonstrated that MLM approaches are superior in the analysis of change or other indirect methods to assess predictors of change. MLM seems to be particularly useful in the context of the present study because participants were not measured at identical time intervals and MLM accommodates such unbalanced data reasonably well (Lindenberger & Ghisletta, 2004).

The MLM approach uses a particular form of the generalized linear model to estimate fixed effects and random effects. In its simplest form, a first-level equation of MLM models individual levels and patterns of development on a given variable as a function of change over time (on a given metric such as time in study). In analogy to latent change or latent growth curve models (McArdle & Nesselroade, 1994), MLM proposes an average initial level (fixed intercept) and a common time-based trajectory for the sample (fixed slope). Individual levels and trajectories are represented as deviations from the sample average. Hereby, these models allow to estimate the variance of an intercept (random intercept) and of a slope factor (random slope) as separate parameters, which, in turn, provides the possibility to determine whether or not there are reliable individual differences in intercept and change over time. Having rejected the null hypothesis that no variance exists is the first and basic step in examining factors that might play a role in individual (or group) differences. A second-level equation can then model individual differences in the above regression coefficients as a function of individual-difference variables that are introduced as covariates.

In the present study, intraindividual stability and change over time as well as interindividual differences therein were examined by estimating individual growth modeling, as implemented by SAS (1997) Proc Mixed,[6] for cognitive functioning, personality and self-related functioning, and social integration. The general procedure was as follows. In a first step, the hypothesis was tested that there were significant individual differences in intraindividual

[6] SAS Proc Mixed utilizes an unstructured error covariance matrix (representing within-person residuals) which imposes no structure on the data except heteroscedasticity (correlated intercepts and slopes). Although an autoregressive pattern may yield a better fit to the data (containing off-diagonal elements that indicate decreasing correlations over time), in many applications an autoregressive pattern is no longer needed, once other fixed and random effects were taken into account (J. Singer, 1998).

change on the four composite measures of psychological functioning. In a second step, it was examined whether individual differences in change trajectories could be explained by selected predictor variables through applying a multilevel growth model with covariates. Using the standard procedure (e.g., Mroczek & Spiro, 2003), each model yielded estimates of fixed effects, which describe the intercept (M_L) and slope (M_S) of the overall sample, and of random effects, which describe the person level in terms of their deviations in intercepts (V_L) and slope (V_S) estimates (Rogosa, 1995).

To represent covariates of differences in intercept and slope over time, interaction terms were used as recommended by J. Singer (1998). Gender, life-history factors, subgroup membership, and age were utilized as covariates. Gender was used as a dummy-coded variable with men coded 0 and women coded 1. Life-history factors were represented by a unit-weighted composite of occupational prestige, years of education, and income. Based on previous analyses (i.e., repeated measures ANOVAs, Schaie procedure), subgroup membership was represented by a dummy-coded variable that contrasted the group that was found to exhibit differential change against both other subgroups combined. For example, the Disparate-Profile Subgroup was contrasted on cognitive functioning against both other subgroups pooled and the Overall-Positive Profile Subgroup was contrasted on self and personality, social integration, and the overall functioning/desirability measure against both other subgroups. To be in parallel to other BASE analyses that have used some kind of individual growth modeling techniques (e.g., T. Singer et al., 2003; Ghisletta & Lindenberger, 2003), age was centered at the mean of the total cross-sectional BASE sample (85). Centering was done to reduce the correlation between intercept and slope that otherwise would be inflated (Kreft, de Leeuw, & Aiken, 1995; Rogosa & Willett, 1985; Willett, 1988).[7]

It was decided to model change over 'time in study' rather than chronological age. Selecting this time dimension was done because the focus was on interindividual differences in intraindividual change. Had the focus been on age difference in change over time, one would have studied change trajectories over chronological age. In the context of the present study, the age span modeled would have been large (i.e., 34 years) relative to the mean longitudinal observation period (6 years). In addition, estimates in the higher age ranges would have been less stable due to small cell sizes above age 85. Finally, modeling change over time in study allows to include age as a covariate to examine whether age played a role in determining change trajectories.

[7] It has to be acknowledged, though, that centering has also been criticized (e.g., Kromrey & Foster-Johnson, 1998).

Each domain-specific composite measure was regressed on linear functions of time in study. For each measure, a sequence of increasingly complex (nested) models were tested and compared against one another on the basis of their model fit. Once a model that allowed individuals to vary in both level and rate of change from the overall trajectory had been found superior, covariates were added to assess their impact. As suggested in the literature (Bryk & Raudenbush, 1992; Mroczek & Spiro, 2003), models that allowed random effects only for the intercept (intercept-only models) were used to estimate the intra-class correlation that indicates the amount of within- and between-person variance. For cognition, the intra-class correlation was .80 suggesting that 80% of the total variation in cognition was between-person variance, and the remainder (20%) was within-person variation. Between-person variation accounted for the majority of variability, but there also was variability within persons suggesting that there was a reasonable amount of individual differences in change over time. The figures on the other three measures were comparable (self and personality: 69%; social integration: 68%, and overall functioning/desirability: 77%).

Multilevel Models for Cognitive Functioning. Table C.3 shows results from the sequence of increasingly complex Multilevel Models for cognitive functioning. The intercept (M_L) represents the predicted mean on a given measure of psychological functioning at baseline assessment, the slope (M_S) is the predicted amount of change over the observation period on that measure. The fixed effects define the sample level. The random effect estimates indicate individual differences relative to the sample level. If the variance of the random effect of the intercept (V_L) is significant, then interindividual differences in mean level exist. If the variance of the random effect of the slope (V_S) is significant, this indicates interindividual differences in intraindividual change over six years.

Model 0 in Table C.3 is the a simple null model that does not allow individuals to differ from one another. Model 1 represents the above mentioned intercept-only model, which includes an intercept (M_L) and variance around the intercept (V_L). In this case, V_L is statistically significant indicating that BASE participants differ in their level of cognitive functioning. Model 4 allowed individuals to vary in both level and rate of linear change. As compared with previous models (0-3), more parameters have to be estimated and the gain in fit (*-2LL*) is substantive. This suggests that there were indications that individuals differ in both mean levels and rates of change over time so that covariates can be introduced to account for the differences. Accordingly, the subsequent models in Table C.3 add different combinations of the covariates. For example, Model 6 that includes a highly significant interaction between the Disparate-Profile Subgroup and change over time was found to show a superior model fit.

Across successive models that included different combinations of the covariates, Model 15 appeared to represent the data best. This model included the following characteristics: (a) fixed and random effects for the intercept (M_L, V_L), (b) no average change over time in cognitive functioning (M_S), but individual differences in change (V_S) although they were not significant in the presence of the covariates, (c) significant effects on mean levels of cognitive functioning for membership in the Overall-Positive Profile Subgroup (M_{L*Sub1}), Age (M_{L*Age}), Gender (M_{L*Sex}), and life-history factors (M_{L*SES}), and (d) significant effects on change over time by membership in the Disparate-Profile Subgroup (M_{S*Sub3}).

Table C. 3
Multilevel Models of Cognitive Functioning: Model Comparison

Model	# more parms[z]	*-2LL*
0 M_L***	2	2,604
1 M_L***, V_L***	3	2,310
2 M_L***, V_L***, M_s***	4	2,292
3 M_L***, V_L***, V_s*	4	2,299
4 M_L***, V_L***, M_s***, $V_s^{\dagger}$, $r_{L,S}$	6	2,285
5 M_L***, V_L***, M_s***, $V_s^{\dagger}$, M_{L*Sub3}, $r_{L,S}$	7	2,283
6 M_L*, V_L***, M_s^a , V_s, M_{L*Sub3}, M_{s*Sub3}***, $r_{L,S}$**	**8**	**2,273**
7 M_L***, V_L***, M_s**, V_s, M_{L*Age}**, M_{S*Age}**, $r_{L,S}$	8	2,274
8 M_L***, V_L***, M_s***, V_s, M_{L*Sex}, M_{S*sex}, $r_{L,S}$	8	2,282
9 M_L***, V_L***, M_s, V_s, M_{L*SES}***, M_{S*SES}, $r_{L,S}$	8	2,282
10 M_L***, V_L***, M_s^a, V_s, M_{L*Sub3}, M_{s*Sub3}***, M_{L*Age}***, M_{L*Sex}**, M_{L*SES}***, $r_{L,S}$	11	2,239
11 M_L***, V_L***, M_s*, V_s, M_{L*Sub3}, M_{s*Sub3}**, M_{S*Age}**, $M_{S*Sex}^{\dagger}$, M_{S*SES}, $r_{L,S}$	11	2,279
12 M_L***, V_L***, M_s, V_s, M_{L*Sub3}, M_{s*Sub3}**, M_{L*Age}***, M_{L*Sex}**, M_{L*SES}***, M_{S*Age}*, M_{S*Sex}, M_{S*SES}, $r_{L,S}$	14	2,250
13 M_L***, V_L***, M_s*, V_s, M_{s*Sub3}***, M_{L*Age}***, M_{L*Sex}***, M_{L*SES}***, $r_{L,S}$	10	2,242
14 M_L***, V_L***, M_s*, V_s, M_{L*Sub1}***, M_{s*Sub3}**, M_{L*Age}***, M_{L*Sex}***, M_{L*SES}***, $r_{L,S}$	11	2,228
15 M_L*, V_L***, V_s, M_{L*Sub1}***, M_{s*Sub3}**, M_{L*Age}**, M_{L*Sex}***, M_{L*SES}***, $r_{L,S}$**	**10**	**2,230**

Note. Superior models in bold. [z] – Number of additional parameters, M_L – Fixed effect for intercept, M_S – Fixed Effect for slope, V_L – Random effect for intercept, V_S – Random effect for slope, $r_{L,S}$ – Covariance of intercept, slope.
SES - Unit-weighted composite of occupational prestige, years of education, and income, Sub 3 - Disparate-Profile Subgroup contrasted against the Overall-Positive Profile and the Average-Profile Subgroup combined, Sub 1 - Overall-Positive Profile Subgroup contrasted against the Average-Profile and the Disparate-Profile Subgroup combined.
* $p < .05$, ** $p < .01$, *** $p < .001$, † $p = .06$, ^{a} $p = .07$.

Table C.4 shows the fixed and random effect estimates from Model 15. The fixed effects are shown in the top half of the Table and random effects are shown in the bottom half. Note that for cognitive functioning, the intercept was significant indicating that the average level of cognitive performance was 57 in a *T* score metric which is about two thirds of a standard deviation above the mean of the total cross-sectional BASE sample. It has to be kept in mind, however, that the intercept was identified using Multilevel Modeling, whereas the mean is directly based on observed data. In addition, the intercept was adjusted for age, gender, life-history factors, and membership in the Overall-Positive Profile Subgroup. The estimate for age was negative indicating that older participants showed lower levels of cognitive performance at baseline assessment. The dummy-coded variable for gender represented men as 0 and women as 1 so that the positive estimate indicates that women were cognitively fitter than men by a factor of 3.3. Similarly, members of the Overall-Positive Profile Subgroup showed better cognitive functioning than members of both other subgroups by a factor of 3.5. It was also found that higher status on life-history factors such as education, income, and professional prestige was linked to better cognitive performance. In line with results obtained from the repeated measures ANOVAs and from applying Schaie's procedure (1989, 1990), the negative estimate of –.55 for the interaction term of the Disparate-Profile Subgroup by change slope over time indicates that members of this subgroup showed stronger decline trajectories on cognitive functioning as compared to both other subgroups pooled. The random effects, fit statistics such as *–2LL* and *AIC*, the sample size, and the number of observations (3 occasions for each of the 130 participants) are also shown.

Table C. 4
Final Multilevel Model for Cognitive Functioning, Including Age, Gender, SES and Subgroup Membership

Fixed Effects	Estimates (*se*)	*t* (*df*)
Intercept	57.47 (6.80)	t (125) = 8.45***
Age	– 0.23 (0.08)	t (125) = – 2.92**
Gender	3.28 (0.94)	t (125) = 3.49***
Overall-Positive Profile Subgroup	3.46 (0.98)	t (125) = 3.54***
SES	0.26 (0.06)	t (125) = 4.45***
Slope x Disparate-Profile Subgroup	– 0.55 (0.12)	t (259) = – 4.82***
Random Effects	Estimates (se)	*z*
Variance of Intercept	21.36 (3.71)	5.76***
Variance of Slope	0.63 (0.43)	1.49
Covariance of Intercept, Slope	0.07 (0.08)	0.91
Residual Variance	7.77 (0.90)	8.66***
-2 LL	2230	
AIC	–1119	
N	130	
Observations	390	

Note. SES: Unit-weighted composite of occupational prestige, years of education, and income. Standard errors in parentheses.
* $p < .05$, ** $p < .01$, *** $p < .001$.

Comparing the random effect size estimates between models 4 and 5 (see J. Singer, 1998) revealed that subgroup membership accounted for 44% of the individual differences in mean level of cognitive performance, and 45% of the individual differences in change over time were explained by subgroup; this is a substantial effect size.

Multilevel Models for Personality and Self-Related Functioning. From Table C.5 the comparison coefficients for models of self and personality can be obtained. Again, a model was found superior that allowed individuals to vary in intercept and slope (4). Introducing the Overall-Positive Profile Subgroup as a covariate for both intercept and slope revealed a significant estimates for both and Model 6, which contained these covariates, fitted better than the models specified before. Model 14 was found superior and it contained the following characteristics: (a) fixed and random effects for the intercept (M_L, V_L), (b) no change over time in self and personality, on average, but individual differences in change (V_S) although they were not significant in the presence of the covariates, (c) significant effects on mean levels of self and personality by membership in the Overall-Positive Profile Subgroup (M_{L*Sub1}) as well as in the Disparate-Profile Subgroup (M_{L*Sub3}), and (d) significant effects on change over time by membership in the Overall-Positive Profile Subgroup (M_{S*Sub1}). Models that included any of the other covariates revealed a worse fit to the data.

Table C. 5
Multilevel Models of Self and Personality: Model Comparison

Model	# more parms[z]	*-2LL*
0 M_L***	2	2,310
1 M_L***, V_L***	3	2,118
2 M_L***, V_L***, M_s***	4	2,094
3 M_L***, V_L***, V_s**	4	2,105
4 M_L***, V_L***, M_s***, V_s*, $r_{L,S}$	6	2,088
5 M_L***, V_L***, M_s***, V_s*, M_{L*Sub1}***, $r_{L,S}$	7	2,044
6 M_L*, V_L***, M_s, V_s, M_{L*Sub1}***, M_{s*Sub1}***, $r_{L,S}$**	**8**	**2,032**
7 M_L***, V_L***, M_s, V_s*, M_{L*Age}**, M_{S*Age}, $r_{L,S}$	8	2,086
8 M_L***, V_L***, M_s**, V_s*, M_{L*Sex}, M_{S*sex}, $r_{L,S}$	8	2,088
9 M_L***, V_L***, M_s, V_s*, M_{L*SES}, M_{S*SES}, $r_{L,S}$	8	2,098
10 M_L***, V_L***, M_s, V_s, M_{L*Sub1}***, M_{s*Sub1}***, M_{L*Age}*, M_{L*Sex}, M_{L*SES}, $r_{L,S}$	11	2,033
11 M_L***, V_L***, M_s, V_s, M_{L*Sub1}***, M_{s*Sub1}***, M_{S*Age}*, M_{S*Sex}, M_{S*SES}, $r_{L,S}$	11	2,045
12 M_L***, V_L***, M_s, V_s, M_{L*Sub1}***, M_{s*Sub1}***, M_{L*Age}, M_{L*Sex}, M_{L*SES}, M_{S*Age}, M_{S*Sex}, M_{S*SES}, $r_{L,S}$	14	2,049
13 M_L***, V_L***, V_s, M_{L*Sub1}***, M_{s*Sub1}***, $r_{L,S}$	7	2,029
14 M_L*, V_L***, V_s, M_{L*Sub1}***, M_{s*Sub1}***, M_{L*Sub3}*, $r_{L,S}$**	**8**	**2,022**
15 M_L***, V_L***, V_s, M_{L*Sub1}***, M_{s*Sub1}***, M_{L*Sub3}*, M_{s*Sub3}, $r_{L,S}$	9	2,024

Note. Superior models in bold. [z] – Number of additional parameters, M_L – Fixed effect for intercept, M_S – Fixed Effect for slope, V_L – Random effect for intercept, V_S – Random effect for slope, $r_{L,S}$ – Covariance of intercept, slope. SES - Unit-weighted composite of occupational prestige, years of education, and income, Sub 3 - Disparate-Profile Subgroup contrasted against the Overall-Positive Profile and the Average-Profile Subgroup combined, Sub 1 - Overall-Positive Profile Subgroup contrasted against the Average-Profile and the Disparate-Profile Subgroup combined.
* $p < .05$, ** $p < .01$, *** $p < .001$.

Table C.6 shows the fixed and random effect estimates from Model 14 from Table C.5. It can be obtained that the average level of self-related functioning and personality was 50 in a *T* score metric, which also was the mean of the total cross-sectional sample. Again, it has to be acknowledged, however, that the intercept has been adjusted for membership in the subgroups, and that the intercept was derived from Multilevel Modeling rather than being observed directly. Members of the Overall-Positive Profile Subgroup were functioning better with regard to self and personality by a factor of 4.3, whereas members of the Disparate-Profile Subgroup were lower in terms of functioning/desirability by a factor of 2.2. Consistent with previous results, the Overall-Positive Profile Subgroup declined on measures of self and personality, as indicated by a negative and statistically significant estimate for the interaction with slope. Subgroup membership accounted for 43% of the individual differences in mean level and 33% of the individual differences in change over time. Again, both effect sizes were substantive.

Table C. 6
Final Multilevel Model for Self and Personality, Including Membership in Overall-Positive Profile Subgroup and Disparate-Profile Subgroup

Fixed Effects	Estimates (*se*)	*t* (*df*)
Intercept	50.42 (0.65)	t (127) = 77.12***
Overall-Positive Profile Subgroup	4.33 (0.80)	t (127) = 5.40***
Disparate-Profile Subgroup	– 2.15 (0.85)	t (127) = – 2.54*
Slope x Overall-Positive Profile Subgroup	– 0.52 (0.08)	t (259) = – 6.34***
Random Effects	Estimates (*se*)	*z*
Variance of Intercept	8.67 (1.75)	4.96***
Variance of Slope	0.25 (0.24)	1.01
Covariance of Intercept, Slope	0.10 (0.06)	1.65
Residual Variance	4.92 (0.60)	8.22***
-2 LL	2022	
AIC	–1015	
N	130	
Observations	390	

Note. Standard errors in parentheses.
* *p* < .05, ** *p* < .01, *** *p* < .001.

Multilevel Models for Social Integration. Table C.7 shows the model comparison coefficients for social integration. Again, Model 4 that included random effects for intercept and slope fitted better than the models before, and an inclusion of the Overall-Positive Profile Subgroup as a covariate of intercept and slope also increased the model fit substantively (see Model 6). Model 15 was the final superior model and included significant effects of membership in the Overall-Positive Profile Subgroup (M_{L*Sub1}), in the Disparate-Profile Subgroup (M_{L*Sub3}), and age (M_{L*Age}) on the intercept. The interaction effect between the Overall-Positive Profile Subgroup and change over time was also found to be significant (M_{S*Sub1}).

Table C. 7
Multilevel Models of Social Integration: Model Comparison

Model	# more parms[z]	*-2LL*
0 M_L***	2	2,582
1 M_L***, V_L***	3	2,397
2 M_L***, V_L***, M_s***	4	2,385
3 M_L***, V_L***, V_s**	4	2,386
4 M_L***, V_L***, M_s***, V_s*, $r_{L,S}$*	6	2377
5 M_L***, V_L***, M_s***, V_s*, M_{L*Sub1}***, $r_{L,S}$	7	2314
6 M_L*, V_L***, M_s, V_s*, M_{L*Sub1}***, $M_{s,Sub1}$***, $r_{L,S}$**	**8**	**2,303**
7 M_L***, V_L***, M_s, V_s*, M_{L*Age}**, M_{S*Age}, $r_{L,S}$*	8	2,369
8 M_L***, V_L***, M_s**, V_s*, M_{L*Sex}*, M_{S*sex}, $r_{L,S}$*	8	2,372
9 M_L***, V_L***, M_s, V_s*, M_{L*SES}, M_{S*SES}, $r_{L,S}$*	8	2,385
10 M_L***, V_L***, M_s, V_s*, M_{L*Sub1}***, M_{s*Sub1}***, M_{L*Age}*, M_{L*Sex}*, M_{L*SES}, $r_{L,S}$	11	2,297
11 M_L***, V_L***, M_s*, $V_s^{\dagger}$, M_{L*Sub1}***, M_{s*Sub1}***, M_{S*Age}**, M_{S*Sex}, M_{S*SES}, $r_{L,S}$	11	2,308
12 M_L***, V_L***, M_s, V_s, M_{L*Sub1}***, M_{s*Sub1}***, M_{L*Age}, M_{L*Sex}*, M_{L*SES}, M_{S*Age}*, M_{S*Sex}, M_{S*SES}*, $r_{L,S}$	14	2,304
13 M_L***, V_L***, V_s*, M_{L*Sub1}***, M_{s*Sub1}***, M_{L*Age}*, M_{L*Sex}*, $r_{L,S}$	9	2,290
14 M_L***, V_L***, V_s*, M_{L*Sub1}***, M_{s*Sub1}***, M_{L*Age}*, M_{L*Sex}, M_{L*Sub3}***, $r_{L,S}$	10	2,273
15 M_L*, V_L***, V_s^a, M_{L*Sub1}***, M_{s*Sub1}***, M_{L*Age}*, M_{L*Sub3}***, $r_{L,S}$**	**9**	**2,277**
16 M_L***, V_L***, V_s*, M_{L*Sub1}***, M_{s*Sub1}***, M_{L*Age}*, M_{L*Sex}, M_{L*Sub3}***, M_{s*Sub3}, $r_{L,S}$	11	2,274

Note. Superior models in bold. [z] – Number of additional parameters, M_L – Fixed effect for intercept, M_S – Fixed Effect for slope, V_L – Random effect for intercept, V_S – Random effect for slope, $r_{L,S}$ – Covariance of intercept, slope.
SES - Unit-weighted composite of occupational prestige, years of education, and income, Sub 3 - Disparate-Profile Subgroup contrasted against the Overall-Positive Profile and the Average-Profile Subgroup combined, Sub 1 - Overall-Positive Profile Subgroup contrasted against the Average-Profile and the Disparate-Profile Subgroup combined.
* $p < .05$, ** $p < .01$, *** $p < .001$, † $p = .06$, ^{a} $p = .07$.

Table C.8 shows fixed and random effects from the superior Model 15 on social integration from Table C.7. The intercept was 61 in a *T* score metric, which is more than one standard deviation above the mean of the total cross-sectional BASE sample. Again, one has to keep in mind that statistics derived from Multilevel Modeling are not directly comparable to those derived from observed constructs and that adjustment were made. Age was found to be negatively associated with baseline status of social integration. In line with both other sets of analyses, the Overall-Positive Profile Subgroup was socially embedded (5.9), whereas there was a clear lack of social integration for members of the Disparate-Profile Subgroup (–4.7). Finally, members of the Overall-Positive Profile Subgroup showed stronger decline on measures of social integration over the six years of observation than did both other subgroups combined. 53% of the individual differences in mean level and 23% of the individual differences in slopes on social integration were accounted for by the subgroups.

Table C. 8
Final Multilevel Model for Social Integration, Including Age, Membership in Overall-Positive Profile Subgroup and Disparate-Profile Subgroup

Fixed Effects	Estimates (*se*)	*t* (*df*)
Intercept	61.06 (5.19)	t (126) = 11.76***
Age	– 0.13 (0.07)	t (126) = – 1.99*
Overall-Positive Profile Subgroup	5.88 (1.01)	t (126) = 5.82***
Disparate-Profile Subgroup	– 4.68 (1.05)	t (126) = – 4.45***
Slope x Overall-Positive Profile Subgroup	– 0.63 (0.12)	t (259) = – 5.18***
Random Effects	Estimates (*se*)	*z*
Variance of Intercept	12.68 (3.05)	4.16***
Variance of Slope	0.03 (0.49)	0.06
Covariance of Intercept, Slope	0.26 (0.13)	1.96
Residual Variance	10.44 (1.27)	8.24***
-2 LL	2277	
AIC	–1143	
N	130	
Observations	390	

Note. Standard errors in parentheses.
* $p < .05$, ** $p < .01$, *** $p < .001$.

Multilevel Models for Overall Functioning/Desirability. From Table C.9 can be obtained that Model 4 that included random effects for intercept and slopes fitted much better than the previously specified models. Using the Overall-Positive Profile Subgroup as a covariate for intercept and slope again improved the model fit (see Model 6), and the final and superior Model 13 included the Overall-Positive Profile Subgroup (M_{L*Sub1}, M_{S*Sub1}) and age (M_{L*Age}, M_{S*Age}) as covariates of both intercept and change slopes.

Table C. 9
Multilevel Models of Overall Functioning/Desirability: Model Comparison

Model	# more parms[z]	*-2LL*
0 M_L***	2	2,223
1 M_L***, V_L***	3	1,964
2 M_L***, V_L***, M_s***	4	1,918
3 M_L***, V_L***, V_s**	4	1,947
4 M_L***, V_L***, M_s***, $V_s^{\dagger}$, $r_{L,S}$	6	1,913
5 M_L***, V_L***, M_s***, $V_s^{\dagger}$, M_{L*Sub1}***, $r_{L,S}$*	7	1,809
6 M_L*, V_L***, M_s*, V_s, M_{L*Sub1}***, M_{s*Sub1}***, $r_{L,S}$****	**8**	**1,797**
7 M_L***, V_L***, M_s, $V_s^{\dagger}$, M_{L*Age}***, M_{S*Age}, $r_{L,S}$	8	1,892
8 M_L***, V_L***, M_s***, $V_s^{\dagger}$, M_{L*Sex}, M_{S*sex}, $r_{L,S}$	8	1,915
9 M_L***, V_L***, M_s, $V_s^{\dagger}$, M_{L*SES}*, M_{S*SES}, $r_{L,S}$	8	1,918
10 M_L***, V_L***, M_s*, V_s, M_{L*Sub1}***, M_{s*Sub1}***, M_{L*Age}***, M_{L*Sex}, M_{L*SES}*, $r_{L,S}$	11	1,787
11 M_L***, V_L***, M_s**, V_s, M_{L*Sub1}***, M_{s*Sub1}***, M_{S*Age}***, M_{S*Sex}, M_{S*SES}*, $r_{L,S}$	11	1,798
12 M_L***, V_L***, M_s*, V_s, M_{L*Sub1}***, M_{s*Sub1}***, M_{L*Age}**, M_{L*Sex}, M_{L*SES}, M_{S*Age}***, M_{S*Sex}, M_{S*SES}, $r_{L,S}$	14	1,795
13 M_L*, V_L***, M_s**, V_s, M_{L*Sub1}***, M_{s*Sub1}***, M_{L*Age}**, M_{S*Age}**, $r_{L,S}$***	**10**	**1,786**

Note. Superior models in bold. [z] – Number of additional parameters, M_L – Fixed effect for intercept, M_S – Fixed Effect for slope, V_L – Random effect for intercept, V_S – Random effect for slope, $r_{L,S}$ – Covariance of intercept, slope.
SES - Unit-weighted composite of occupational prestige, years of education, and income, Sub 3 - Disparate-Profile Subgroup contrasted against the Overall-Positive Profile and the Average-Profile Subgroup combined, Sub 1 - Overall-Positive Profile Subgroup contrasted against the Average-Profile and the Disparate-Profile Subgroup combined.
* $p < .05$, ** $p < .01$, *** $p < .001$, † $p = .06$.

Finally, Table C.10 shows estimates gained from the final Model 13 in examining differences and change over time in the overall measure of functioning/desirability. The intercept is again relatively high (59 in a *T* score metric) and the sample was found to show even an increase in functioning (1.64). Of course, this is a function of covarying out the influence of age and the Overall-Positive Profile Subgroup. Age was associated with lower baseline levels (–.11) and stronger decline over time (–.02). Subgroup membership was linked to higher levels of baseline functioning/desirability (6.00) and also stronger decline over time (–.40). At the zero-order level, for example, an intercept of 51.4 and a significant decline of –.25 were found. The subgroup variable was found to explain 70% of the individual differences in mean level and 40% of the individual differences in change over time on the overall measure of functioning/desirability. Both effect sizes can be regarded substantial.

Table C. 10
Final Multilevel Model for Overall Functioning/Desirability, Including Age and Membership in Overall-Positive Profile Subgroup

Fixed Effects	Estimates (*se*)	*t* (*df*)
Intercept	59.22 (3.09)	t (127) = 19.18***
Slope	1.64 (0.55)	t (257) = 2.97**
Age	– 0.11 (0.04)	t (127) = – 2.97**
Overall-Positive Profile Subgroup	6.00 (0.46)	t (127) = 13.02***
Slope x Age	– 0.02 (0.01)	t (257) = – 3.23**
Slope x Overall-Positive Profile Subgroup	– 0.40 (0.08)	t (257) = – 4.88***
Random Effects	Estimates (*se*)	*z*
Variance of Intercept	3.79 (0.86)	4.41***
Variance of Slope	0.24 (0.12)	1.98*
Covariance of Intercept, Slope	0.03 (0.03)	0.88
Residual Variance	2.82 (0.34)	8.41***
-2 LL	1786	
AIC	– 897	
N	130	
Observations	390	

Note. Standard errors in parentheses.
* $p < .05$, ** $p < .01$, *** $p < .001$.

To sum up the above results: Regression towards the mean is an issue threatening each and every attempt to examine change over time on a given set of measures. The approach taken in the present study is *one* way to deal with this fundamental question. Having found convergent evidence across different methods suggests that it is rather unlikely that subgroup differences in change over time can completely be accounted for by regression artifacts. Instead, differential stability and decline over a period of six years in BASE appears to represent a robust and reliable phenomenon.

C. 2. 2 The Role of External Correlates For Subgroup Change Over Time

Table C.11 supplements the age and gender analyses are reported in Section 4.2.3 and provides descriptive information for repeated measures ANOVAs using change over six years on a given psychological measure as within-subject factor (3 occasions) and Subgroup membership at baseline assessment (3 groups), Age Cohort (2 groups; cut-off: 85 years at the third occasion), and Gender as between-subject factors. It can be obtained that oldest-old participants from BASE showed lower levels of functioning than did the young old, whereas there were virtually no differences between men and women.

To be inserted

Table C. 11
Change Over Time on the Four Composite Measures of Psychological Functioning by Age Cohort (85 years at Third Wave) and Gender

Multiple Regression Analyses. Table C.12 supplements multiple regression analyses reported in Section 4.2.3 that used subgroup membership, physical functioning, and life-history factors to predict change over time. The Table reports means, standard deviations, and zero-order Pearson correlation coefficients for the four predictor variables as well as the four outcome variables. Bivariate correlations, for example, revealed that change in overall functioning/desirability was negatively associated with the Overall-Positive Profile Subgroup (*r* = –.32) and positively with the Disparate-Profile Subgroup (*r* = .23). No bivariate associations were found for both the physical-functioning factor and the life-history factor. In addition, some intercorrelations among the predictor variables were sizeable. First, the dummy-coded variables for membership in the Overall-Positive Profile and the Disparate-Profile Subgroup correlated, of course, negatively (*r* = –.64). Second, subgroups were significantly, but differentially associated with physical functioning (Overall-Positive Profile: *r* = .37; Disparate Profile: *r* = – .23), suggesting that participants in the Overall-Positive Profile Subgroup were functioning better and those from the Disparate-Profile Subgroup were functioning worse. Third, life-history factors showed no associations to the grouping variables, but the intercorrelation with physical functioning was significant (*r* = .26).

Table C. 12
Means, Standard Deviations, and Intercorrelations for Predictor and Outcome Variables Used in Multiple Regression Analyses of Subgroup Status, Physical Functioning, and Life-History Factors to Predict Change over 6 Years in the Four Composite Measures of Psychological Functioning

			Intercorrelation			
Measure	*M*	SD	1	2	3	4
Predictors						
1. Overall-Positive Subgroup	.47	*.50*	–			
2. Disparate-Profile Subgroup	.32	*.47*	**–.64**	–		
3. Physical Functioning	.00	*1.00*	**.37**	**–.23**	–	–
4. Life-History Factors	.00	*1.00*	.17	.13	**.26**	–
Outcomes						
5. Δ Overall Desirability	–.27	*.50*	**–.32**	**.23**	.05	.02
6. Δ Cognition	–.26	*.77*	.11	**–.30**	.14	–.07
7. Δ Self & Personality	–.27	*.69*	**–.30**	**.25**	–.05	–.03
8. Δ Social Integration	–.29	*1.02*	**–.32**	**.36**	.03	.13

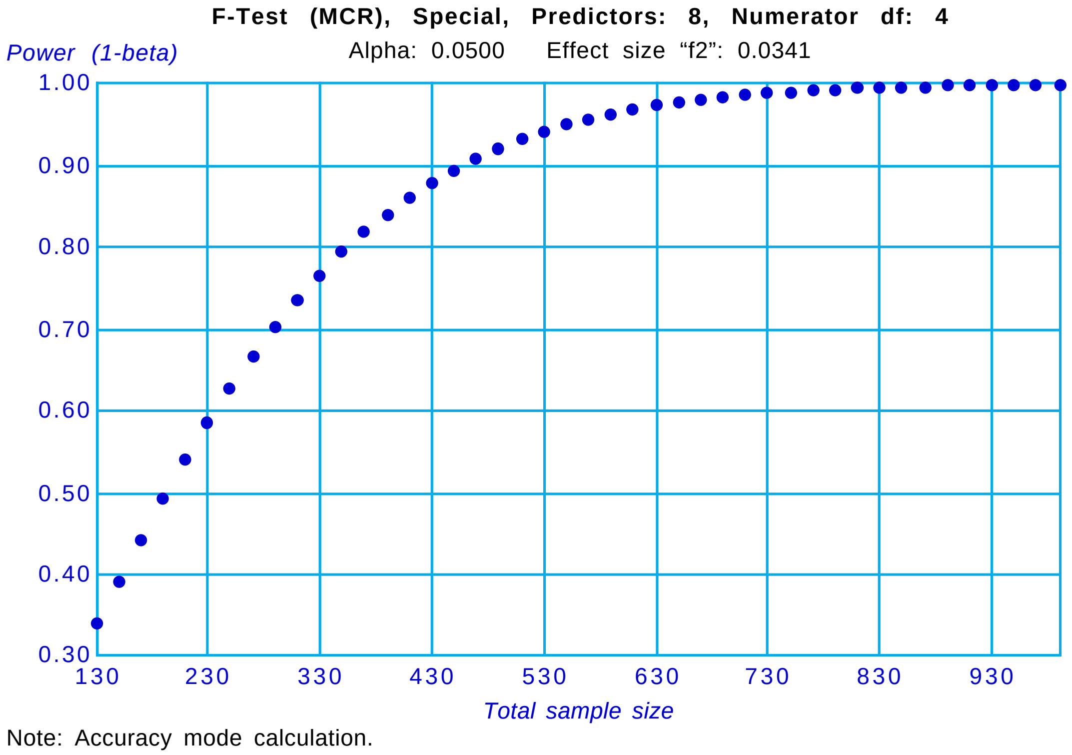
Note. *N* = 130; Physical Functioning: Unit-weighted composite of multimorbidity and sensory functioning; Life-history factors: Unit-weighted composite of occupational prestige, years of education, and income. Coefficients in bold are significantly different from zero at the *p* < .05 level or below.

Power Analyses. Restrictions in statistical power may be one central limiting factor in finding statistically significant effects. Power analyses (carried out using G*Power) indicated that, given the small sample size, statistical power to determine a small interaction effect (e.g., .03 unique variance out of .12 overall variance explained; Cohen, 1977) was very low (*power* = .034). A sample size of about *n* = 450 participants would have been necessary to assure sufficient statistical power for detecting such effect size (*power* ≥ 90; see Table C.13 for Panel A). Given the present sample size (*n* = 130), statistical power to determine interaction effects would have been reasonably high only if the interaction effect itself were of medium size (e.g., effect size ≥ .12; *power* ≥ .95; see Table C.13 for Panel B).

Table C. 13
Summary of Power Analyses using G*Power

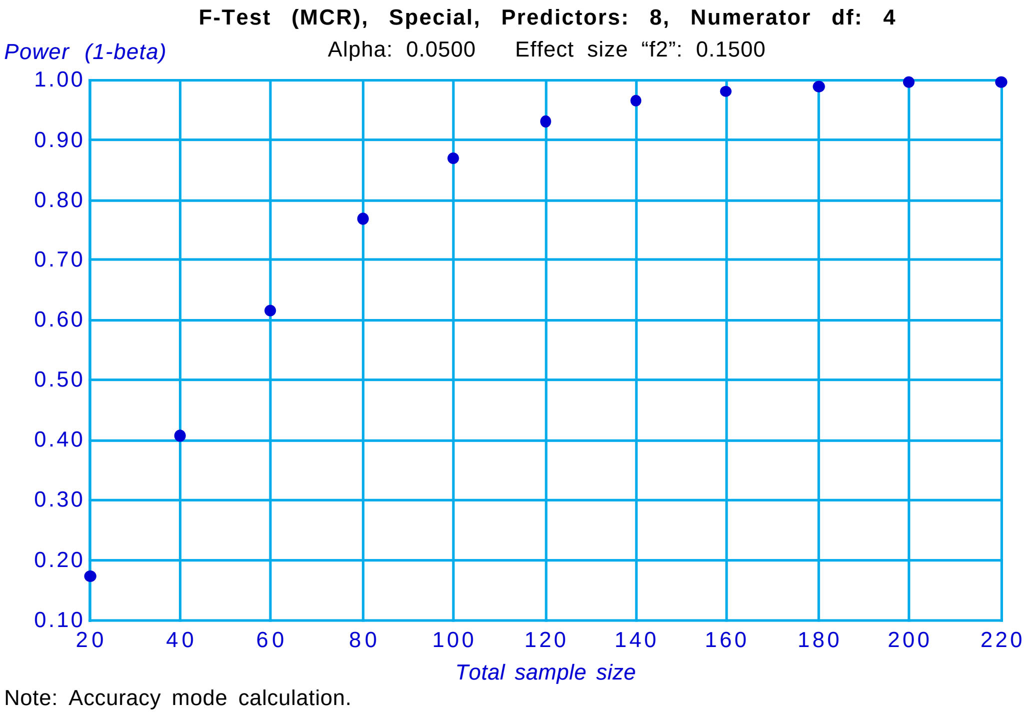
(A) Post-hoc analysis for "F-Test (MCR)", Special, Predictors: 8, Numerator df: 4:

Alpha: 0.0500,	Power (1-beta): 0.3402
Effect size "f2": 0.0341	Total sample size: 130
Critical value: F(4,121) = 2.4466	Lambda: 4.4330



(B) Post-hoc analysis for "F-Test (MCR)", Special, Predictors: 8, Numerator df: 4:

Alpha: 0.0500
Power (1-beta): 0.9520
Effect size "f2": 0.1500
Total sample size: 130
Critical value: F(4,121) = 2.4466
Lambda: 19.5000



C. 3 Outcomes of Heterogeneity and Differential Development Over Time: Consequences of Subgroup Differences in Psychological Profiles and Change Over Six Years

Table C.14 gives means, standard deviations, and intercorrelations for the predictor and outcome variables used in multiple regression analyses to predict well-being (see Section 4.3.1). It can be seen that the numerically strongest relationship of both well-being measures were found to the grouping variable (r = .34 at both occasions), of course, with the exception of the autocorrelation (r = .60). Similarly strong associations were found between well-being at baseline assessment and multimorbidity (r = .31) as well as between well-being at the last occasion and sensory functioning (r = .30), although the latter association was not found using the baseline measure of well-being (r = .11, p > .10). Interestingly, age, gender, and life-history variables showed numerically weaker associations with the outcome measure of well-being at the third wave than did subgroup membership.

Table C. 14
Means, Standard Deviations, and Intercorrelations for Predictor and Outcome Variables Used in Multiple Regression Analyses of Subgroup Status, Perceptual Speed, Age, Gender, Life-History Factors, Multimorbidity, Sensory Functioning, and Baseline Well-Being to Predict Well-Being Six Years Later

			Intercorrelation							
Measure	*M*	*SD*	1	2	3	4	5	6	7	8
Predictors										
1. Subgroups	.47	.50	–							
2. Perceptual Speed	57.64	7.43	**.38**	–						
3. Age	78.29	5.96	**–.26**	**–.30**	–					
4. Gender	1.55	.50	–.04	.06	.04	–				
5. Life-History Factors	51.95	8.36	.17	**.24**	.06	**–.20**	–			
6. Multimorbidity[a]	53.13	9.34	**.24**	**.26**	–.13	–.16	**.18**	–		
7. Sensory Functioning	56.93	9.07	**.32**	**.52**	**–.48**	.07	**.22**	.17	–	
8. Well-Being at Baseline	51.84	9.26	**.34**	**.18**	–.05	**–.25**	.09	**.31**	.11	–
Outcome										
9. Well-Being at 3rd Wave	49.01	9.95	**.34**	**.18**	**–.19**	–.13	**.18**	**.26**	**.30**	**.60**

Note. N = 130; Life-history factors: Unit-weighted composite of occupational prestige, years of education, and income. [a] Scores on this dimension were recoded to calculate the desirability (functional status) score. Coefficients in bold are significantly different from zero at the p < .05 level or below.

APPENDIX D DISCUSSION

D. 1 Gender Differences in the Total Cross-Sectional and the 6-Year Longitudinal BASE Sample

Table D.1 and Table D.2 report gender differences in the total cross-sectional BASE sample (*N* = 516; Table D.1) and in the 6-year longitudinal BASE sample (*N* = 130; Table D.2). Comparing both Tables with one another demonstrates that the overwhelming majority of differences between men and women that were evident in the total cross-sectional sample were not present when restricting the analyses to those participants who took part three times in BASE. Specifically, 24 out of 30 constructs revealed statistically significant gender differences in the cross-sectional sample. In contrast, only 6 out of 30 measures were found to significantly differ between men and women in the longitudinal BASE sample.

…

Table D. 1
Comparison of the Life Contexts of Men (N = 258) and Women (N = 258) in the Total Cross-Sectional BASE Sample: Social Status, Life-History Factors, Physical Health, and Intellectual Functioning

Measure	Men	Women	Significant effects for gender
Social status and life history			
% Married	52	7	$\chi^2(3, N = 516) = 127.9, p < .001$
% Widowed	40	70	
% Divorced	4	11	
% Single	4	12	
Mean no. of children	1.46	1.05	$F(1, 515) = 18.9, p < .001$
Mean no. of living kin	4.7	3.6	$F(1, 513) = 6.7, p < .01$[b]
% Living alone	38	65	$\chi^2(1, N = 516) = 94.0, p < .001$
% Institutionalized	10	18	$\chi^2(1, N = 516) = 6.7, p < .001$
Social participation (T score)[a]	50.4	49.6	*ns*
Mean household income (DM)	2,208[c]	1,877	$F(1, 515) = 13.3, p < .001$
Education (years)	11.3	10.2	$F(1, 513) = 28.2, p < .001$
Occupational prestige (T score)	51.2	48.9	$F(1, 515) = 6.5, p < .05$
Physical health			
Multimorbidity (T score)	48.7	51.1	$F(1, 515) = 18.9, p < .01$
Cardiovascular illness (T score)	49.7	50.3	*ns*
Muscular-skeletal illness	48.9	51.9	$F(1, 515) = 19.7, p < .001$
Other illness (e.g., cancer)	49.5	50.5	*ns*
Subjective health (5-point scale)	3.06	2.75	$F(1, 515) = 11.0, p < .001$
Functional capacity			
Sensory (composite)	50.1	49.9	*ns*
Visual acuity (T score)	51.2	48.8	$F(1, 513) = 10.2, p < .01$
Hearing acuity (T score)	49.1	50.8	$F(1, 513) = 7.0, p < .01$
ADL and IADL[d] (T score)	51.7	48.2	F(1, 513) = 22.8, p < .001
Gait and balance (T score)	50.9	49.1	$F(1, 504) = 8.9, p < .01$
			Age x Gender, $F(1, 504) = 2.4, p < .05$
Mental health			
HRSD (T score)	48.3	51.7	$F(1, 515) = 15.1, p < .001$
CES-D (T score)	48.1	51.8	$F(1, 504) = 17.3, p < .001$
% Diagnosed as depressive	20	31	$\chi^2(1, N = 516) = 8.1, p < .01$
MMSE	51.1	48.5	F(1, 515) = 10.1, $p < .01$
			Age x Gender, $F(1, 504) = 3.9, p < .05$
% Suspected dementia[c]	17	25	$\chi^2(1, N = 516) = 4.1, p < .05$
Intellectual ability			
g factor (composite score)	50.8	49.1	$F(1, 515) = 4.0, p < .05$
Perceptual speed	50.6	49.4	*ns*
Reasoning	51.4	48.6	$F(1, 515) = 12.8, p < .001$
Memory	49.4	50.6	*ns*
Fluency	50.9	49.1	$F(1, 515) = 4.2, p < .05$[f]
Knowledge	51.6	48.4	$F(1, 515) = 14.4, p < .000$

Note. DM = Deutsche Mark; ADL = Activities of Daily Living; IADL = Instrumental Activities of Daily Living; HRSD = Hamilton Rating Scale for Depression; CES-D = Center for Epidemiological Studies – Depression Scale; MMSE = Mini Mental State Exam; BASE = Berlin Aging Study; DSM-III-R = Diagnostic and Statistical Manual of Mental Disorders (3rd. Ed., Rev.). [a] When the original scale is non informative, means are reported as T scores (M = 50, SD = 10).
[b] Degree-of-freedom values other than (1, 515) indicate missing values. [c] Per capita income weighted according to size of household. Amounts in U.S. dollars are approximately $1,300 versus $1,000. [d] IADL items known to be biased by gender roles (e.g., meal preparation, light housework) were excluded. [e] In BASE, 109 individuals were diagnosed as having suspected dementia (Helmchen et al., 1999). The diagnoses were differentiated at three levels: Level 1 (n = 37) indicated participants with symptomatic memory difficulties, and Level 2 and 3 (ns = 33 ad 39, respectively) indicated cases with moderate-to-severe symptomatology (DSM-III-R). In all of these individuals, language functioning was still intact. They were able to respond to questions and o produce sentences. For research ethical reasons, individuals with severe dementia that involved the loss of language were excluded from participation in the study. [f] Note that Lindenberger and Baltes (1997) did not report that gender was significant for fluency. Their analyses set probability values at Bonferroni-adjusted levels because the abilities are correlated. For information (and to be consistent with the other analyses described in this table), we report the unadjusted values here.
(adapted from Smith & Baltes, 1998).

Table D. 2

Comparison of the Life Contexts of Men (N = 59) and Women (N = 71) in the 6-Year Longitudinal BASE Sample: Social Status, Life-History Factors, Physical Health, and Intellectual Functioning

Measure	Men	Women	Significant effects for gender
Social status and life history			
% Married	38	6	$\chi^2(3, N = 130) = 49.2, p < .001$
% Widowed	17	49	
% Divorced	0	11	
% Single	4	5	
Mean no. of children	0.75	0.72	*ns*
Mean no. of living kin	5.1	4.0	*ns*
% Living alone	21	59	$\chi^2(1, N = 130) = 30.7, p < .001$
% Institutionalized	1	2	*ns*
Social participation (T score)[a]	56.2	58.0	*ns*
Mean household income (DM)	2,553[c]	1,969	$F(1, 129) = 4.9, p < .05$
Education (years)	54.8	49.9	$F(1, 129) = 7.8, p < .01$
Occupational prestige (T score)	51.9	52.1	*ns*
Physical health			
Multimorbidity (T score)	54.8	51.7	*ns*
Cardiovascular illness (T score)	54.5	53.6	*ns*
Muscular-skeletal illness	49.5	49.5	*ns*
Other illness (e.g., cancer)	54.4	49.5	$F(1, 129) = 9.3, p < .01$
Subjective health (5-point scale)	53.1	52.3	*ns*
Functional capacity			
Sensory (composite)	55.4	56.4	*ns*
Visual acuity (T score)	57.2	54.8	*ns*
Hearing acuity (T score)	53.5	58.0	$F(1, 129) = 6.4, p < .05$
ADL and IADL[d] (T score)	56.2	55.2	*ns*
Gait and balance (T score)	0.5	0.5	*ns*
Mental health			
HRSD (T score)	52.5	50.5	*ns*
CES-D (T score)	10.4	12.6	*ns*
% Diagnosed as depressive	10	15	*ns*
MMSE	53.7	54.4	*ns*
% Suspected dementia[c]	1	2	*ns*
Intellectual ability			
g factor (composite score)	57.1	58.4	*ns*
Perceptual speed	56.8	58.3	*ns*
Reasoning	57.8	56.0	*ns*
Memory	53.9	58.6	$F(1,129) = 9.8, p < .01$
Fluency	55.8	58.1	*ns*
Knowledge	56.4	55.2	*ns*

Note. DM = Deutsche Mark; ADL = Activities of Daily Living; IADL = Instrumental Activities of Daily Living; HRSD = Hamilton Rating Scale for Depression; CES-D = Center for Epidemiological Studies – Depression Scale; MMSE = Mini Mental State Exam; BASE = Berlin Aging Study; DSM-III-R = Diagnostic and Statistical Manual of Mental Disorders (3rd. Ed., Rev.).

[a] When the original scale is non informative, means are reported as T scores (M = 50, SD = 10).

[b] Degree-of-freedom values other than (1,129) indicate missing values.

[c] Per capita income weighted according to size of household. Amounts in U.S. dollars are approximately $1,300 versus $1,000.

[d] IADL items known to be biased by gender roles (e.g., meal preparation, light housework) were excluded.

[e] In the BASE, 109 individuals were diagnosed as having suspected dementia (Helmchen et al., 1999). The diagnoses were differentiated at three levels: Level 1 (n = 37) indicated participants with symptomatic memory difficulties, and Level 2 and 3 (ns = 33 ad 39, respectively) indicated cases with moderate-to-severe symptomatology (DSM-III-R). In all of these individuals, language functioning was still intact. They were able to respond to questions and o produce sentences. For research ethical reasons, individuals with severe dementia that involved the loss of language were excluded from participation in the study.

[f] Note that Lindenberger and Baltes (1997) did not report that gender was significant for fluency. Their analyses set probability values at Bonferroni-adjusted levels because the abilities are correlated. For information (and to be consistent with the other analyses described in this table), we report the unadjusted values here.

Lebenslauf

13.06.1976	Geboren in Quedlinburg
1983 – 1991	Besuch der POS ‚Karl Liebknecht' Hoym
1991 – 1995	Besuch des Gymnasiums ‚Askaneum' Aschersleben
1995	Abitur am Gymnasium ‚Askaneum' Aschersleben
1997 – 2001	Studium der Psychologie an der Freien Universität Berlin
April 1999	Vordiplom in Psychologie
1999 – 2001	Studentische Hilfskraft am Max-Planck-Institut für Bildungsforschung, Berlin
September 2001	Diplom in Psychologie
2001 – 2004	Promotionsstipendium am Max-Planck-Institut für Bildungsforschung, Berlin

Erklärung

Hiermit versichere ich, dass ich die vorgelegte Arbeit selbständig verfasst habe.

Andere als die angegebenen Hilfsmittel habe ich nicht verwendet.

Die Arbeit ist in keinem früheren Promotionsverfahren angenommen oder abgelehnt worden.

Berlin, den 12. 10. 2004

Denis Gerstorf