

Empfehlungen der NAKO- Expertengruppe „Neurologisch- Psychiatrische Erkrankungen“ für die Datenanalyse der NAKO Gesundheitsstudie

Recommendations of the
NAKO expert group “Neurologic-
Psychiatric Diseases” for data
analyses within the German
National Cohort [NAKO]

[VERSION 1, NOV. 19, 2025]

Empfehlungen der NAKO-Expertengruppe „Neurologisch-Psychiatrische Erkrankungen“ für die Datenanalyse der NAKO Gesundheitsstudie

Recommendations of the NAKO expert group ‚Neurologic-Psychiatric Diseases‘ for data analyses within the German National Cohort (NAKO)

(Version 1, nov. 19, 2025)

- 1) Empfehlungen zum Umgang mit der Neuropsychologischen Testbatterie
[Recommendations for handling results of the neuropsychological test battery](#)
- 2) Empfehlungen zum Umgang mit dem Fragenset ‚Subjektive Gedächtnisprobleme‘
[Recommendations for handling the questionnaire “Subjective Memory Problems”](#)
- 3) Empfehlungen zur Klassifikation der Testergebnisse in 3 Kategorien abnehmender Kognitionsleistung
[Recommendations for the classification of test results into three categories reflecting decreasing levels of cognitive performance](#)
- 4) Empfehlungen zur Klassifikation einer Parkinsonerkrankung (Baseline) anhand von Selbstangaben und Medikation
[Recommendations for the baseline classification of Parkinson’s disease based on self-reported information and medication data](#)
- 5) Empfehlungen zum Umgang mit Sniffin Sticks Ergebnissen für Stift 6 und Papierstreifen
[Recommendations for handling Sniffin’ Sticks results for stick 6 and paper strips](#)

Anmerkung: Alle Empfehlungen gelten für die NAKO Erst- und Zweituntersuchung und die aufgeführten Variablen [*kursiv*] sind im NAKO Transfer Hub für beide Untersuchungszeitpunkte verfügbar. Für Fragen zu den Empfehlungen kontaktieren Sie bitte die Expertengruppenkoordinatoren: Klaus Berger, Universität Münster (bergerk@uni-muenster.de) oder Fabian Streit, Zentralinstitut für Seelische Gesundheit Mannheim (fabian.streit@zi-mannheim.de)

Note: All recommendations are applicable to the NAKO baseline (1U) and first follow-up (2U) examinations and all listed derived variables [*italic*] are available for both datasets. For questions please contact the expert group coordinators: Klaus Berger, University of Münster (bergerk@uni-muenster.de) or Fabian Streit, Central Institute of Mental Health Mannheim (fabian.streit@zi-mannheim.de)

1) Datenanalyse der neuropsychologischen Testbatterie

Data analyses using the neuropsychological test battery

A) Zu beantragende Variablen

Um eine umfassende Auswertung der neuropsychologischen Testbatterie aus Level 1 und Ausschluss möglicher Störfaktoren zu ermöglichen, sollten mindestens die folgenden, abgeleiteten Variablen im NAKO Transfer Hub beantragt werden:

1. Tiernamen [*a_npsy_sumts1*]
2. Wortliste 1-3 [*a_npsy_sumts2_1*, *a_npsy_sumts2_2*, *a_npsy_sumts2_3*]
3. Stroop Tafel 1, 2, Tafel 2-1 [*a_npsy_sumts3_2*, *a_npsy_sumts3_3*, *a_npsy_diffs3_32*]
4. Digit Span [*a_npsy_sumts4*]
5. Störfaktoren [*a_npsy_stoer_seh*, *a_npsy_stoer_hoer*, *a_npsy_stoer_orga*, *a_npsy_stoer_indiv*, *a_npsy_stoer_abbruch*] (s. Punkt C)
6. Sprachfähigkeit [*a_ses_deutsch*] (s. Punkt D)

Die genannten Tests und ihre Durchführung sind in der folgenden Publikation beschreiben, auf die als Referenz verwiesen werden kann: Kleineidam L., Stark M, Riedel-Heller SG, Pabst A et al: World J Biol Psychiatry. 10(2023), 909-923, <https://doi.org/10.1080/15622975.2021.2011408>

Zusätzlich sind der Perdue-PegBoard-Test, ein Test der Koordination und Feinmotorik und der Zahlenreihen-Test, ein Maß der fluiden Intelligenz, nur bei Level 2 NAKO-Probanden erhoben worden. Der Zahlenreihen-Test ist in der folgenden Publikation, auf die als Referenz verwiesen werden kann, ausführlich beschrieben: Schmiedek F., Kroehne U, Goldhammer F, Prindle JJ et al: World J Biol Psychiatry. 10(2023), 924-935, <https://doi.org/10.1080/15622975.2021.2011407>

Des Weiteren sollten immer Alter [*basis_age*], Geschlecht [*basis_sex*] und Bildungsjahre [*a_ses_isced97_years*] oder Bildungslevel [*a_ses_isced11_level*] oder Schulabschluss [*d_se_b1*] oder alle drei beantragt werden.

To enable a comprehensive evaluation of the neuropsychological test battery from Level 1 and to rule out potential interfering factors, at minimum the following, derived variables should be applied for in the NAKO Transfer Hub:

- Animal names [*a_npsy_sumts1*]
- Word list 1-3 [*a_npsy_sumts2_1*, *a_npsy_sumts2_2*, *a_npsy_sumts2_3*]
- Stroop board 1, 2, board 2-1 [*a_npsy_sumts3_2*, *a_npsy_sumts3_3*, *a_npsy_diffs3_32*]
- Digit Span [*a_npsy_sumts4*]
- Interfering factors [*a_npsy_stoer_seh*, *a_npsy_stoer_hoer*, *a_npsy_stoer_orga*, *a_npsy_stoer_indiv*, *a_npsy_stoer_motor*, *a_npsy_stoer_abbruch*] (see item C)
- Language proficiency [*a_ses_deutsch*] (see item D)

The single tests and their administration are described in detail in the following publication, which can be referenced: Kleineidam L., Stark M, Riedel-Heller SG, Pabst A et al: World J Biol Psychiatry. 10(2023), 909-923, <https://doi.org/10.1080/15622975.2021.2011408>

Additionally, the Purdue PegBoard Test, a test of coordination and fine motor skills, and the Number-Series-Test, a measure of fluid intelligence, have been administered in Level 2 participants. The Number-Series-Test is described in detail in the following publication, which can

be referenced: Schmiedek F., Kroehne U, Goldhammer F, Prindle JJ et al: World J Biol Psychiatry. 10(2023), 924-935, <https://doi.org/10.1080/15622975.2021.2011407>

Finally, it is recommended to always apply for age [*basis_age*], gender [*basis_sex*], and years of education [*a_ses_isced97_years*] or level of education [*a_ses_isced11_level*] or school qualification [*d_se_b1*] or for all three.

B) Test 1: Tiernamen (Animal names, semantic memory)

Die Variable „Anzahl falsch genannte Tiernamen“ [*npsy_ts1_2*] sollte nicht verwendet werden, da sie unvollständig (nicht bei allen Teilnehmenden, nicht in allen Studienzentren) erhoben wurde.

The variable ‚Number of incorrectly named animals‘ (Anzahl falsch genannter Tiernamen [*npsy_ts1_2*]) should not be used because of its incomplete assessment.

C) Umgang mit den Störvariablen (How to deal with ‚interference variables‘)

Die während der neuropsychologischen Testung aufgetretenen und von Study Nurses dokumentierten Störungen wurden im Rahmen der Qualitätssicherung und Plausibilisierung in 6 Störkategorien plus einem Stör-Summenscore eingeteilt. Zu den Störfaktoren zählen:

- Organisatorische Störfaktoren (Lärm von draußen, technischer Defekt etc. [*a_npsy_stoer_organ*])
- Abbruch eines Tests (*a_npsy_stoer_abbruch*)
- Individuelle Störfaktoren (Schlafmangel, Nachtschicht, unmotiviert etc. [*a_npsy_stoer_indiv*])
- Hörprobleme (Schwerhörig, Hörgerät vergessen, taub, etc. [*a_npsy_stoer_hoer*])
- Sehprobleme (Brille vergessen, blind etc. [*a_npsy_stoer_seh*])
- Motorische Einschränkungen (Fingeramputation, rheumatol. Arthritis, Arm in Gips etc. [*a_npsy_stoer_motor*])
- Summe der einzelnen Störfaktoren (*a_npsy_stoersum*)

Störfaktoren können die Durchführung einzelner oder mehrerer Tests unmittelbar beeinträchtigen, sodass die erhobenen Testwerte die tatsächliche kognitive Leistungsfähigkeit der Probanden nicht mehr länger adäquat abbilden. In diesem Fall ist ein Ausschluss dieser Probanden/innen von der Analyse sinnvoll, um eine systematische Fehleinschätzung der Kognition bei bestimmten Probandengruppen zu vermeiden. Störfaktoren können aber auch Proxies für andere Faktoren sein z.B. den allgemeinen Gesundheitszustand, in diesem Fall ist ein Ausschluss nicht sinnvoll. Der Störfaktoren-Summenscore bietet eine einfache Möglichkeit multiple potentielle Störungen zu erkennen.

Aufgrund des unten in Figure 1 gezeigten Einflusses der einzelnen Störfaktoren auf die Ergebnisse der Einzeltests (gemessen mit Cohen's d, adjustiert für Alter/Geschlecht) werden folgende Empfehlungen gegeben:

- *Probanden, die positiv für die Störvariablen „technisch-organisatorisches Problem“ oder „individuelles Problem“ oder „Testabbruch“ sind, sollten von Auswertungen der neuropsychologischen Testbatterie grundsätzlich ausgeschlossen werden (d.h. Ausschluss von max. 1555 Probanden [0,8% aller mit ≥ 1 Test]). Der Effekt dieser 3 Störfaktoren ist bedeutsam ($d = 0,3-0,6$) und hierbei werden externe Einflussgrößen erfasst, die sich nicht auf Personencharakteristika zurückführen lassen.*
- *Probanden, die positiv sind für die Störvariable „Hörproblem“ (N=2440, 1,2%), sollten für die 3 Tests der Wortliste (Tests 2,3 und 6) sowie dem daraus resultierenden Domänenscore „Gedächtnis“ ausgeschlossen werden, da bereits eine rein sensorische Einschränkung des*

Hörvermögens auf Grund der auditiven Darbietung der Wortliste die Validität der Gedächtnismessung vermindern kann.

- Probanden, die positiv sind für die Störvariable „motorisches Problem“ sollten für die Tests der Level 1 Batterie **nicht** ausgeschlossen werden, da für diese lediglich verbale Antworten erforderlich sind und somit durch motorische Einschränkungen keine Testinterferenz stattgefunden hat. Stattdessen gehen die Assoziationen mit der Testleistung auf ein individuelles Merkmal (z.B. Fingeramputation, Rheuma) zurück, das ggfs. Testergebnisse beeinflusst. Probanden mit dieser Störvariable sollten allerdings bei Auswertungen des Level 2 Tests Perdue Peg Board (Test 5) ausgeschlossen werden, da dieser Test eine motorische Antwort erfordert.
- Probanden, die positiv sind für die Störvariable „Sehprobleme“ sollten **nicht** ausgeschlossen werden, da die Assoziationen mit der Testleistung lediglich eine sehr kleine Effektstärke aufweisen.

During neuropsychological testing, disturbances observed and documented by study nurses were categorized, as part of the quality assurance and data validation procedures, into six interference (disturbance) categories plus a composite disturbance score. The disturbance factors include:

- Organizational-technical disturbances (e.g., external noise, technical malfunction, etc. [*a_npsy_stoer_orga*])
- Test interruption (*a_npsy_stoer_abbruch*)
- Individual disturbances (e.g., sleep deprivation, night shift, lack of motivation, etc. [*a_npsy_stoer_indiv*])
- Hearing problems (e.g., hearing impairment, forgot hearing aid, deafness [*a_npsy_stoer_hoer*])
- Visual problems (e.g., forgot glasses, blindness, etc. [*a_npsy_stoer_seh*])
- Motor impairments (e.g., finger amputation, rheumatoid arthritis, arm in cast, etc. [*a_npsy_stoer_motor*])
- Sum of all disturbance factors (*a_npsy_stoersum*)

Interference factors may directly interfere with the administration of one or more tests; in such cases, it is reasonable to exclude these participants from analysis. However, disturbance factors may also serve as proxies for other variables (e.g., general health status); in these cases, exclusion is not appropriate. The summary score of interference factors offers a simple way to detect multiple interferences.

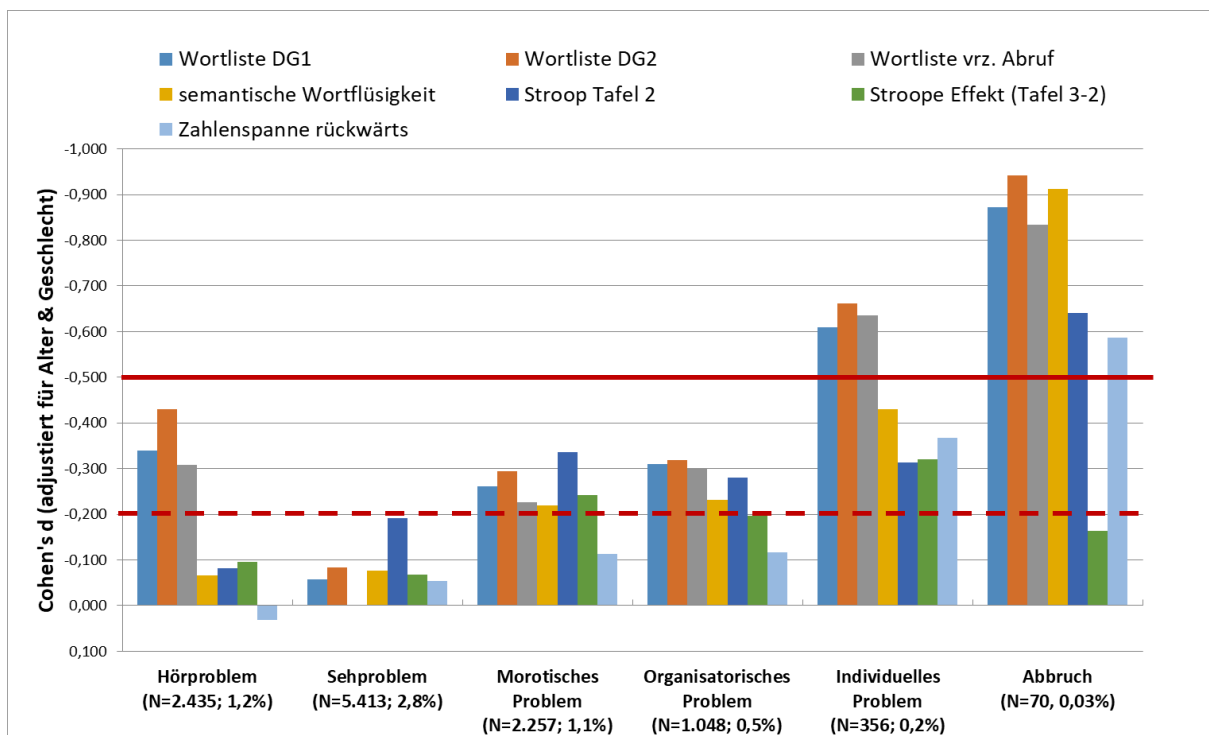
Based on the influence of individual disturbance factors on single-test results, as shown in Figure 1 below (measured by Cohen's *d*, adjusted for age and sex), the following recommendations are made:

- Participants who are positive for the interference variables “organizational-technical problem”, “individual problem” or “test interruption” should generally be excluded from all analyses of the neuropsychological test battery (i.e., exclusion of up to 1,555 participants [0.8% of all with ≥ 1 test]). The effects of these three interference factors are substantial ($d = 0.3-0.6$).
- Participants positive for the interference variable “hearing problem” ($N = 2,440$; 1.2%) should be excluded from analyses of the three word list tests (Tests 2, 3, and 6) and from the derived domain score “Memory,” since even a purely sensory limitation of hearing ability can reduce the validity of the memory assessment due to the auditory presentation of the word list..
- Participants positive for the interference variable “motor problem” should **not** be excluded from analyses of tests of the Level 1 battery, since these tests only require verbal responses, and

therefore no test interference occurred due to motor impairments. Instead, the associations with the test performance are due to an individual characteristic (e.g., finger amputation, rheumatism) that may influence the test results. But participants with this interference variable **should** be excluded from analyses of the Level 2 test Purdue Pegboard (Test 5).

- Participants positive for visual impairments should **not** be excluded, as the associations with test performance have only a very small effect size.

Figure 1: Untersuchung der Assoziation zwischen einzelnen Störfaktoren und Ergebnissen der neuropsychologischen Einzeltests (Examination of the association between individual interference factors and the results of the neuropsychological subtests.)



D) Umgang mit der deutschen Sprachfähigkeit (Handling of German Language Proficiency)

Mehrere der neuropsychologischen Tests weisen eine starke Abhängigkeit von der Kenntnis der deutschen Sprache und der Fähigkeit sie zu sprechen auf. Deswegen ist bei jeder Auswertung der neuropsychologischen Testdaten zu prüfen, inwieweit Assoziationen mit Kognition von der Sprachfähigkeit der Probanden abhängen (durch Adjustierung, Stratifikation, Interaktionsanalyse o.ä.). Dafür wurde im Rahmen der Qualitätssicherung die Variable ‚Sprachfaehigkeit [a_ses_deutsch] neu abgeleitet.

- a_ses_deutsch
- 0 = Deutsche/r Muttersprachler/in
 - 1 = Deutsch plus weitere Muttersprache
 - 2 = Kein Muttersprachler, Deutschkenntnis sehr gut
 - 3 = Kein Mutterspr., Deutschkenntnis gut
 - 4 = Kein Mutterspr., Deutschkenntnis mittelmäßig
 - 5 = Kein Mutterspr., Deutschkenntnis schlecht
 - 6 = Kein Mutterspr., Deutschkenntnis sehr schlecht

Several neuropsychological tests are highly dependent on knowledge of the German language and the ability to speak it. Therefore, whenever neuropsychological test data is analysed, it is necessary to examine the extent to which associations with cognition depend on the language skills of the participants (through adjustment, stratification, interaction analysis, etc.). As part of the quality assurance, the variable “*Language Proficiency*” [*a_ses_deutsch*] was newly derived:

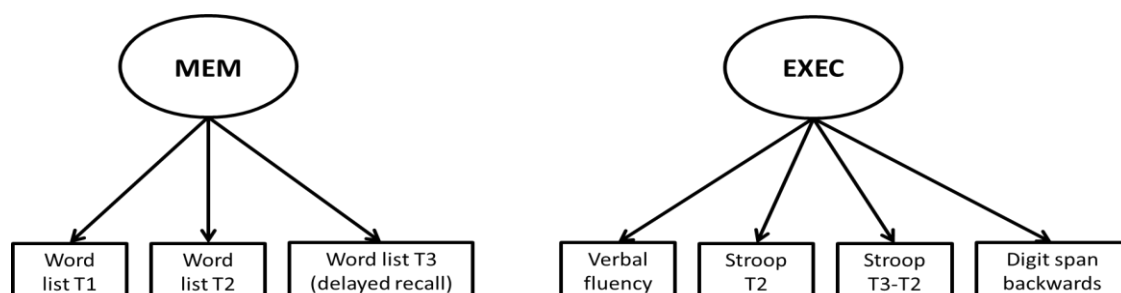
- a_ses_deutsch* 0 = Native German speaker
- 1 = German plus another native language
- 2 = Non-native, very good German proficiency
- 3 = Non-native, good German proficiency
- 4 = Non-native, moderate German proficiency
- 5 = Non-native, poor German proficiency
- 6 = Non-native, very poor German proficiency

E) Zusammenfassen von Testergebnissen - Domänenscores „Gedächtnis“ und „Exekutivfunktion“ (Aggregation of test results – Domain scores “Memory” and “Executive Function”)

Die Einzeltestergebnisse können durch eine Faktorenanalyse in zwei kognitive Domänen, *Gedächtnis* und *Exekutive Funktion*, zusammengefasst werden. Beide Domänenscores wurden bereits abgeleitet und beschrieben (Kleineidam L., Stark M, Riedel-Heller SG, Pabst A et al: *World J Biol Psychiatry*. 10(2023), 909-923, <https://doi.org/10.1080/15622975.2021.2011408>). Die Zuordnung der Einzeltestergebnisse zu den Domänen sind der Figur 2 zu entnehmen. Aus dem Mittelwert beider Domänen ergibt sich ein *Globaler Kognitionsfaktor*. Faktorscores, die auf weniger als 2 Testergebnissen pro Domäne beruhen, sollten auf missing gesetzt werden. Faktorscores sollten z-skaliert (Mittelwert=0, Standardabweichung=1) werden, basierend auf Mittelwert und Standardabweichung aus der Baseline Erhebung. Um eine Rekonstruktion der abgeleiteten Faktorscores zu erlauben, können Link R-Skripte und Modellparameter angefragt werden: transfer@nako.de

The individual test results can be summarized into two cognitive domains, *Memory* and *Executive Function*, using factor analysis. Both domain scores have already been derived and described (Kleineidam L., Stark M., Riedel-Heller S.G., Pabst A. et al.: *World J Biol Psychiatry*, 10(2023), 909–923, <https://doi.org/10.1080/15622975.2021.2011408>). The assignment of individual test results to the domains are shown in Figure 2. The mean of both domain scores yields a *Global cognition factor*. Factor scores should be set to missing if there were fewer than two tests measured per cognitive domain. Factor scores should be z-standardized (mean = 0, standard deviation = 1) based on the mean and standard deviation from the baseline assessment. To enable reconstruction of the derived factor scores, R scripts and model parameters can be requested: transfer@nako.de

Figure 2: Zuordnung der Einzeltests zu den Domänen ‚Gedächtnis‘ und ‚Exekutive Funktion‘ (Assignment of individual tests to the domains “Memory” and “Executive Function.”)



Um die Umsetzung der Empfehlungen zum Umgang mit Störfaktoren und Sprachfähigkeit zu vereinfachen, wurden drei Varianten der Domänenscores abgeleitet:

1. Ohne jeden Probandenausschluss. Diese Scores entsprechen nicht (!) den Empfehlungen unter C) und D). Resultierende Domänennamen: *MEM_Raw, EXEC_Raw, GLOBAL_Raw*.
2. Umsetzung der Empfehlungen zu Störfaktoren (C): Probanden ohne eine oder mehrere der folgenden Störvariablen: ‚Individuelles Problem‘, ‚Organisatorisch-technisches Problem‘ oder ‚Abbruch‘. Dabei wurden für die Domäne ‚Gedächtnis‘ zusätzlich Probanden mit der Störvariable ‚Hörproblem‘ ausgeschlossen. Resultierende Domänennamen: *MEM, EXEC, GLOBAL*.
3. Probanden wie unter 2. beschrieben, aber mit zusätzlichem Ausschluss aller Probanden, die keine Muttersprachler/innen sind (deutsch oder deutsch plus 1 weitere Muttersprache). Diese Scores können zur Umsetzung der Empfehlung (D) genutzt werden. Resultierende Domänennamen: *MEM_NativeSpeaker, EXEC_NativeSpeaker, GLOBAL_NativeSpeaker*.

To simplify the implementation of the recommendations for dealing with interfering factors and language ability, three variants of the domain scores were derived.:

1. No participant exclusions. These scores do not (!) correspond to the recommendations under C) and D). Resulting domain names: *MEM_Raw, EXEC_Raw, GLOBAL_Raw*.
2. Implementation of recommendations on interference variables (C): Excluding participants with one or more of the following interference variables: “Individual Problem,” “Organizational-Technical Problem,” or “Test Interruption.” For the “Memory” domain, participants with the interference variable “Hearing Problem” were additionally excluded. Resulting domain names: *MEM, EXEC, GLOBAL*.
3. Excluding participants as described in 2, with the additional exclusion of all participants who are not native speakers of German (German only or German plus one additional native language). These scores can be used to implement recommendation (D). Resulting domain names: *MEM_NativeSpeaker, EXEC_NativeSpeaker, GLOBAL_NativeSpeaker*.

2) Variablenableitung für den Fragenset ‚Subjektive Gedächtnisprobleme‘ Subjective Memory Complaints. questions in the interview

Im Interview wird an Personen ab einem Alter von 45 Jahre ein Set von fünf Fragen zur subjektiven Wahrnehmung des eigenen Gedächtnisses gestellt. Hintergrund dieser Fragen zum subjektiven Gedächtnis sind Publikationen, die zeigen, dass die Wahrnehmung einer Gedächtnisverschlechterung in Verbindung mit darauf basierender Sorge ein Hinweis für einen rascheren Verlust kognitiver Leistungen sein kann. Im Rahmen der Qualitätssicherung wurden aus dem Fragenset bereits zwei Variablen abgeleitet:

- a_emo_mem5c* 0 = Subjektiv (sehr) gutes Gedächtnis
1 = Subjektiv mässiges Gedächtnis, keine Verschlechterung
2 = Gedächtnis schlechter, keine Sorge, kein Arztkontakt
3 = Gedächtnis schlechter, keine Sorge, mit Arztkontakt
4 = Gedächtnis schlechter mit Sorge, kein Arztkontakt
5 = Gedächtnis schlechter mit Sorge und Arztkontakt
- a_emo_mem2c* 0 = Keine Sorgen wg. des Gedächtnis
1 = Gedächtnis schlechter mit Sorgen darüber

During the interview, a set of five questions regarding the subjective perception of one's own memory is administered to individuals aged 45 years or older. The rationale for these subjective memory questions is based on publications showing that the perception of memory decline, particularly when accompanied by concern, may indicate a more rapid loss of cognitive performance. As part of quality assurance, two variables have already been derived from this question set:

- a_emo_mem5c* 0 = Subjectively (very) good memory
 1 = Subjectively moderate memory, no decline
 2 = Memory worse, no concern, no medical contact
 3 = Memory worse, no concern, with medical contact
 4 = Memory worse with concern, no medical contact
 5 = Memory worse with concern and medical contact
- a_emo_mem2c* 0 = No concerns about memory
 1 = Memory worse with concerns

3) Empfehlungen zur Klassifikation der Testergebnisse in 3 Kategorien abnehmender Kognitionsleistung

Recommendations for the classification of test results into three categories reflecting decreasing levels of cognitive performance

Es werden in der NAKO-Transfer Hub zwei Einteilungen kognitiver Funktion zur Verfügung gestellt, die unterschiedlichen Zwecken dienen sollen

A) Eine Einteilung aller NAKO-Probanden in

1. Unterdurchschnittliche Kognition, d.h. eine kognitive Leistung im globalen Kognitionsfaktor unter dem 10. Perzentil, nach Adjustierung für Alter und Geschlecht
2. Durchschnittliche Kognition, d.h. eine kognitive Leistung im globalen Kognitionsfaktor zwischen dem 10. Perzentil und 90. Perzentil (einschließlich), nach Adjustierung für Alter und Geschlecht
3. Überdurchschnittliche Kognition, d.h. eine kognitive Leistung im globalen Kognitionsfaktor über dem 90. Perzentil, nach Adjustierung für Alter und Geschlecht

Diese Einteilung kann genutzt werden, wenn Personen mit einer ungewöhnlichen hohen oder niedrigen kognitiven Leistungsfähigkeit für ihr Alter und Geschlecht identifiziert oder untersucht werden sollen, unabhängig von möglichen Ursachen.

Die Ableitung der erwarteten Perzentile für ein gegebenes Alter und Geschlecht basiert auf einer Quantil-Regression. Eine R-Syntax zur Ableitung der Grenzwerte und Einteilung in die Kategorien kann angefragt werden: transfer@nako.de.

B) Einteilung der NAKO-Probanden mit deutscher Muttersprache ab 45 Jahre in

1. Normale Kognition
2. Leichtes neuropsychologisches Defizit
3. Schweres neuropsychologisches Defizit

Diese Einteilung kann genutzt werden, wenn es um wahrscheinlich erworbene (durch Krankheiten oder Hirnschädigungen verursachte) kognitive Einschränkungen geht. Neuropsychologische Defizite sind mit höherer Wahrscheinlichkeit erworben, wenn sie deutlich unterhalb des alters- geschlechts- und bildungsadjustierten Erwartungswertes aus einer kognitiv gesunden Normpopulation liegen und im mittleren oder höheren Lebensalter messbar

sind, in dem die Entwicklungsprozesse im Gehirn abgeschlossen sind. Mit steigendem Alter erhöht sich zudem die Prävalenz von Erkrankungen (wie z.B. neurodegenerativen Erkrankungen), die die Hirnfunktion negativ beeinflussen. In der NAKO wurde diese Definition deshalb erst bei Personen ab dem 45. Lebensjahr gebildet und beruht auf einer stärker selektierten Stichprobe zur Ableitung der Grenzwerte. Zur Begründung der Begrenzung auf deutsche Muttersprachler, siehe Punkt 1D.

Da in der NAKO keine Informationen zu kognitiven Veränderungen außerhalb der Gedächtnisdomäne oder einer selbsteingeschätzten Gedächtnisfunktion vorhanden ist und zudem keine Fremdeinschätzung möglicher Alltagsbeeinträchtigungen vorliegt, ist eine Ableitung einer leichten kognitiven Störung (engl. MCI) oder einer Demenz aus den Primärdaten gemäß internationalen Kriterien nicht möglich.

Es wird deswegen empfohlen, auf eine neuropsychologische Definition einer leichten kognitiven Störung anhand des Baserate-Ansatzes zurückzugreifen (Leichtes neuropsychologisches Defizit). Der Baserate-Ansatz (Mistridis, P. et al.: *European Archives of Psychiatry and Clinical Neuroscience* 265:5 (2015): 407-417; <https://doi.org/10.1007/s00406-014-0571-z>) stellt einen etablierten Ansatz zur Definition einer kognitiven Einschränkung dar und zeigt eine hohe prognostische Vorhersagekraft für die Entwicklung einer Demenz (Polcher, A. et al.: *Journal of Alzheimer's Disease* 88:4 (2022): 1663-1678; <https://doi.org/10.3233/JAD-215548>; Oltra-Cucarella, J et al. *Journal of the American Geriatrics Society* 66:7 (2018): 1360:1366. <https://doi.org/10.1111/jgs.15412>). Grenzwerte für die Einteilung in dieser Gruppe beruhen auf einer Quantil-Regression in Personen ab 45 Jahren und keinem Hinweis auf einen Zustand oder Erkrankung, der die kognitive Leistungsfähigkeit beeinträchtigen könnte (Diagnose einer Epilepsie, Parkinson, Schlaganfall, akute depressive Symptomatik [PHQ-9 \geq 10], akut erhöhte Ängstlichkeit [GAD7 \geq 10]). Eine Person weist gemäß dem Baserate-Ansatz eine kognitive Störung auf, wenn diese in mehr kognitiven Tests einer Testbatterie einen auffälligen Wert (d.h. unter dem 16. Perzentil / z-Wert <1 , adjustiert für Alter, Geschlecht und Bildung) aufweist als 10% der Referenzpopulation, die zur Ableitung der Grenzwerte genutzt wurde. In der NAKO war dies bei mehr als 4 (von 10) „auffälligen“ Scores aus der L1 Testbatterie der Fall. Zusätzlich werden Personen als Baserate-MCI klassifiziert, wenn ihre kognitive Leistung im Gedächtnis oder den exekutiven Funktionen (Faktorscores) unter dem 3. Perzentil (adjustiert für Alter, Geschlecht und Bildung) liegt. Dies dient dazu, dass Personen mit einer stark unterdurchschnittlichen Leistung in einer spezifischen kognitiven Domäne nicht als kognitiv normal klassifiziert werden.

Um zwischen leichteren und schwereren kognitiven Störungen, wie man sie etwa im Rahmen einer Demenz erwarten würde, zu unterscheiden, wird zusätzlich empfohlen, Personen mit einem Base-rate MCI und einer kognitiven Leistung im Gedächtnis und dem exekutiven Funktionen Faktor unter dem 3. Perzentil (adjustiert für Alter, Geschlecht und Bildung) als „schweres neuropsychologisches Defizit“ zu klassifizieren. Grenzwerte wurden aus der gleichen Stichprobe gewonnen, die auch zur Entwicklung des leichten neuropsychologischen Defizits genutzt wurde.

Eine R-Syntax zur Ableitung der Grenzwerte und Einteilung in die Kategorien kann angefragt werden: transfer@nako.de.

Two classifications for cognitive function are provided in the NAKO Transfer Hub, serving different purposes:

A) Classification of all NAKO participants into:

1. Below-average cognition: cognitive performance in the global cognition factor below the 10th percentile, after adjustment for age and sex

2. Average cognition: cognitive performance in the global cognition factor between the 10th and 90th percentile (inclusive), after adjustment for age and sex
3. Above-average cognition: cognitive performance in the global cognition factor above the 90th percentile, after adjustment for age and sex

This categorization can be used when individuals with unusually high or low cognitive performance for their age and sex are to be identified or analyzed, regardless of possible causes. The derivation of expected percentiles for a given age and sex is based on quantile regression. An R syntax for deriving the threshold values and assigning participants to categories can be requested: transfer@nako.de.

B) Classification of NAKO participants with German as their native language aged 45 years and older into:

1. Normal cognition
2. Mild neuropsychological deficit
3. Severe neuropsychological deficit

This classification can be used when likely acquired (disease- or brain injury-related) cognitive impairments are of interest. Neuropsychological deficits are more likely to be acquired when their scores are clearly below the age-, sex-, and education-adjusted expected value derived from a cognitively healthy reference population and are measurable in mid-age or older adulthood, when brain developmental processes are complete. With increasing age, the prevalence of diseases (e.g., neurodegenerative disorders) that negatively affect brain function also rises. Therefore, this definition in NAKO was applied only to participants aged 45 years and older and was based on a more selective subsample for deriving the cut-off values. For the rationale behind the restriction to native German speakers, see section 1D.

Since NAKO does not contain information on cognitive changes outside the memory domain, nor on individuals with good self-rated memory function, and since no external assessments of potential everyday impairments are available, it is not possible to derive a diagnosis of Mild Cognitive Impairment (MCI) or dementia from the primary data according to international criteria.

It is therefore recommended to use a neuropsychological definition of mild cognitive impairment based on the base rate approach (“mild neuropsychological deficit”). The base rate approach (Mistridis, P. et al., *European Archives of Psychiatry and Clinical Neuroscience*, 265:5 (2015): 407–417; <https://doi.org/10.1007/s00406-014-0571-z>) is an established method for defining cognitive impairment and shows strong prognostic predictive validity for dementia development (Polcher, A. et al., *Journal of Alzheimer’s Disease*, 88:4 (2022): 1663–1678; <https://doi.org/10.3233/JAD-215548>; Oltra-Cucarella, J. et al., *Journal of the American Geriatrics Society*, 66:7 (2018): 1360–1366; <https://doi.org/10.1111/jgs.15412>). Cut-off values for this classification were derived using quantile regression in participants aged 45 years and older with no indication of a condition or disease that could impair cognitive performance (diagnosis of epilepsy, Parkinson’s disease, stroke, acute depressive symptoms [PHQ-9 \geq 10], or acute anxiety symptoms [GAD-7 \geq 10]). According to the base rate approach, a person is considered to have a cognitive impairment when they show abnormal performance (below the 16th percentile / z-score $<$ -1, adjusted for age, sex, and education) in more cognitive tests of a test battery than 10% of the reference population used to derive the thresholds. In NAKO, this was the case when more than 4 (out of 10) scores from the L1 test battery were below the cut-off. In addition, individuals are classified as having base rate MCI if their cognitive performance in memory or executive function (factor scores) falls below the 3rd percentile (adjusted for age, sex, and education). This ensures that participants with markedly below-average performance in a specific cognitive domain are not classified as cognitively normal. To distinguish between mild and more severe cognitive impairments, such as those typically observed in dementia, it is further recommended that individuals with base rate MCI and cognitive performance in both the

memory and executive function factors below the 3rd percentile (adjusted for age, sex, and education) are classified as having a severe neuropsychological deficit. The cut-off values were derived from the same subsample that was used to define mild neuropsychological deficits.

An R syntax for deriving the threshold values and assigning participants to categories can be requested: transfer@nako.de.

4) Empfehlungen zur Klassifikation einer Parkinsonerkrankung (Baseline) anhand von Selbstangaben und Medikation

Recommendations for the baseline classification of Parkinson's disease based on self-reported information and medication data

Zur Verbesserung der Validität einer Selbstangabe ‚Parkinsonerkrankung‘ wurden die Angaben zu aktuell eingenommenen Medikamenten sowie einer eventuellen Behandlung und der Zeit seit ärztlicher Diagnosestellung herangezogen. Die entsprechenden Variablen sind alle im Baseline Datensatz verfügbar. Tabelle 1 zeigt die daraus basierenden Klassifikationen der Parkinsonerkrankung in sichere oder wahrscheinlich/ mögliche Erkrankung. Diese Klassifikation ist als abgeleitete Variable [*a_diag_park*] bereits verfügbar.

To improve the validity of self-reported 'Parkinson's disease', information on current medication use, any prior treatment, and the time since physician diagnosis was considered. All corresponding variables are available in the baseline dataset. Table 1 shows the resulting classification of Parkinson's disease into definite or probable/possible cases based on these data. This classification is already available as a derived variable [*a_diag_park*].

Table 1: Klassifikation einer Selbstangabe Parkinsonerkrankung (Classification of self-reported Parkinson's disease)

Lifetime physician diagnosis PD	PD treatment last 12 months	Current dopaminergic medication (ATC "N04")	Time since PD diagnosis	Case classification
yes	yes	yes	any	YES definite
yes	yes	no/ unknown	any	YES probable/ possible
yes	no/ unknown	yes	any	YES definite
yes	no/ unknown	no/ unknown	<=5 years	YES probable/ possible
yes	no/ unknown	no/ unknown	>5 years	NO no PD case
don't know	no/ unknown	yes	unknown	YES probable/ possible
no	no	yes	any	NO no PD case

5) Empfehlungen zum Umgang mit Sniffin Sticks Ergebnissen für Stift 6 und Papierstreifen

Recommendations for handling Sniffin' Sticks results for stick 6 and paper strips

Für einen der 12 Stifte (Stift 6) gibt es einen Distraktor (Alternativantwort „Grapefruit“), der inzwischen von einem großen Teil der Probanden anstelle der richtigen Lösung („Zitrone“) genannt wird. Die Gründe für das Anwachsen der Häufigkeit dieser Alternativantwort sind unklar, Veränderungen in der Zusammensetzung der im Stift enthaltenen Flüssigkeit könnten dazu beigetragen haben. Deshalb wurde von den Modulverantwortlichen entschieden für diesen Stift in der NAKO zwei richtige Antworten zu akzeptieren („Zitrone“ und „Grapefruit“). In einer Promotion an der Uni Münster (Stogbauer J et al: *Rhinology*. 2020 Apr 1;58(2):184-191. doi: <https://doi.org/10.4193/Rhin19.181>) konnte gezeigt werden, dass dieses Problem in allen bisherigen deutschen Bevölkerungsstudien mit Sniffin-Sticks-12 Untersuchung auftritt und nicht NAKO spezifisch ist. Die ursprünglich abgeleitete Variable 'olf_rslt', bildet die Summe richtig erkannter Stifte ohne Berücksichtigung der Antwort ‚Grapefruit‘ als richtige Antwort bei Stift 6. Zusätzlich wurde eine neue Variable 'olf_sum' als Summenscore aller Stifte mit Berücksichtigung von 2 richtigen Antworten bei Stift 6 gebildet. Beide Summenscores werden nur gebildet wenn >=10 Stifte getestet wurden. Die weiteren Variablen 'olf_kat' und 'olf_norm' werden jetzt aus 'olf_sum' und nicht mehr aus 'olf_rslt' abgeleitet. Die abgeleiteten Variablen sind für die Datensätze Erst- und Zweituntersuchung verfügbar.

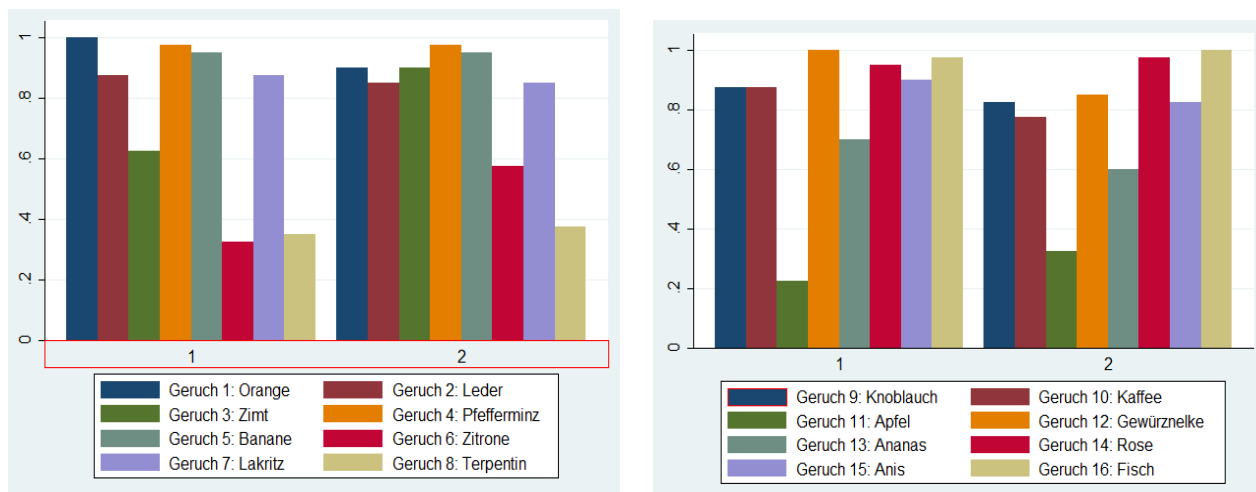
For one of the 12 sticks (stick 6), there is a distractor (alternative response “grapefruit”) that is increasingly reported by participants instead of the correct solution (“lemon”). The reasons for the rising frequency of this alternative response are unclear; changes in the composition of the liquid contained in the stick may have contributed. Therefore, the module coordinators decided to accept two correct answers for this stick in NAKO (“lemon” and “grapefruit”). A doctoral thesis at the University of Münster (Stogbauer J et al.: *Rhinology*, 2020 Apr 1; 58(2):184–191. doi: <https://doi.org/10.4193/Rhin19.181>) demonstrated that this issue occurred in all previous German population studies using the Sniffin-Sticks-12 test and is not specific to NAKO. The originally derived variable *olf_rslt* represents the sum of correctly identified sticks without counting “grapefruit” as a correct answer for stick 6. Additionally, a new variable, *olf_sum*, was created as the total score of all sticks, considering two correct answers for stick 6. Both total scores are calculated only if ≥10 sticks were tested. The variables *olf_kat* and *olf_norm* are now derived from *olf_sum* rather than *olf_rslt*. These derived variables are available for both the baseline and follow-up datasets.

In der Erstuntersuchung war die Sniffin Sticks Untersuchung Teil des längeren L2-Programms. Ursprünglich sollte sie in der Zweituntersuchung nur für ehemalige L2 Teilnehmende wiederholt werden. Aufgrund der Corona Pandemie und des verhängten Lockdowns mussten die NAKO Studienzentren 2020 während der laufenden Zweituntersuchung für mehrere Monate geschlossen werden. Aufgrund der Bedeutung des Symptoms Geruchsverlust bei einer Corona-Infektion wurde die Untersuchung danach aber für alle ins Programm genommen. Aufgrund eines möglichen Infektionsrisikos wurde diese Untersuchungsmethode allerdings in den Jahren 2020-2021 auf Papierstreifen umgestellt. Grundlage für diese Umstellung war ein Vergleich der Geruchserkennung zwischen Stift und Papierstreifen, der im Rahmen einer Masterarbeit an der Universität Münster zuvor durchgeführt wurde. Figur 3 zeigt den Vergleich zwischen Papierstreifen und Stift für 16 verschiedene Stifte. Die Erkennungsleistung für die Mehrheit der Stifte unterschied sich zwischen beiden Erkennungsmethoden nur minimal. Für einige wenige Stifte waren jedoch deutlichere Unterschiede vorhanden. Deswegen wurde im Rahmen der Qualitätssicherung der Zweituntersuchungsdaten eine neue, abgeleitete Variable ‚Sniffin Sticks Methode‘ [*olf_method*] erstellt. Sie sollte bei Analysen der Riechtestergebnisse aus der Zweituntersuchung eingesetzt werden. Auch haben nach Ende des Lockdowns verschiedene Studienzentren die Untersuchung

mit den Sniffin Sticks nur verzögert wieder aufgenommen. Daraus resultieren unterschiedliche Probandenzahlen in den Studienzentren für die Zweituntersuchung.

In the baseline examination, the Sniffin-Sticks test was part of the longer L2 program. Originally, it was intended to be repeated in the follow-up assessment only for former L2 participants. Due to the COVID-19 pandemic and the resulting lockdown, NAKO study centers had to close for several months in 2020 during the ongoing follow-up assessments. Given the significance of the symptom of olfactory loss in COVID-19 infections, the test was subsequently included in the program for all participants. However, due to potential infection risk, this testing method was switched to paper strips in 2020–2021. This change was based on a comparison of odor detection between sticks and paper strips, previously conducted as part of a Master's thesis at the University of Münster. Figure 3 shows the comparison between paper strips and sticks for 16 (1-8 and 9-16) different odors. Detection performance differed only minimally for the majority of odors between the two methods, although some odors showed more pronounced differences. Therefore, as part of quality assurance for the follow-up data, a new derived variable, “*Sniffin-Sticks Method*” [*olf_method*], was created. This variable should be used in analyses of olfactory test results from the first follow-up (2U). Additionally, after the end of the lockdown, some study centers resumed the Sniffin-Sticks testing with delay, resulting in varying participant numbers across centers for the follow-up assessment.

Figure 3: Vergleich der Erkennung von 16 Stiften des erweiterten Sniffin Sticks Sets zwischen Papierstreifen und direktem Riechen am Stift (Comparison of odor identification for 16 items of the extended Sniffin-Sticks set using paper strips versus direct sniffing of the sticks)



Dieses Dokument wurde von der Expertengruppe Neurologische und Psychiatrische Erkrankungen der NAKO Gesundheitsstudie erstellt und ist ausschließlich für den internen Gebrauch im Rahmen der Datennutzung vorgesehen. Eine Weitergabe, Vervielfältigung oder Verwendung außerhalb dieses Zwecks – ganz oder teilweise – ist ohne vorherige Zustimmung der Expertengruppe oder der NAKO- Geschäftsstelle nicht gestattet. Bei Fragen zur Weitergabe oder Nutzung wenden Sie sich bitte an die verantwortliche Expertengruppe: www.nako.de/expertengruppen.

This document was produced by the Expert Group on Neurological and Psychiatric Disorders of the German National Cohort (NAKO) and is intended solely for internal use in the context of data utilisation. Any disclosure, reproduction or use outside the scope of this purpose – in whole or in part – is not permitted without the prior consent of the expert group or the NAKO central office. If you have any questions regarding disclosure or use, please contact the relevant expert group: www.nako.de/expertgroups.