Journal of open psychology data

Data From the MySWOW Proof-of-Concept Study: Linking Individual Semantic Networks and Cognitive Performance

DATA PAPER

DIRK U. WULFF
SAMUEL AESCHBACH
SIMON DE DEYNE
RUI MATA
*Author affiliations can be found in the back matter of this article

ubiquity press

ABSTRACT

We report data from a proof-of-concept study involving the concurrent assessment of large-scale individual semantic networks and cognitive performance. The data include 10,800 free associations—collected using a dedicated web-based platform over the course of several weeks—and responses to several cognitive tasks, including verbal fluency, episodic memory, associative recall tasks, from four younger and four older native German speakers. The data are unique in scope and composition and shed light on individual and age-related differences in mental representations and their role in cognitive performance across the lifespan.

CORRESPONDING AUTHOR:
Dirk U. Wulff
Department of Psychology, University of Basel, Missionsstrasse 60-62, 4055 Basel, Switzerland
dirk.wulff@gmail.com

KEYWORDS:
semantic networks; cognitive aging; individual differences

TO CITE THIS ARTICLE:
Wulff, D. U., Aeschbach, S., De Deyne, S., & Mata, R. (2022). Data From the MySWOW Proof-of-Concept Study: Linking Individual Semantic Networks and Cognitive Performance. *Journal of Open Psychology Data,* 10: 5, pp. 1–8
DOI: *https://doi.org/10.5334/jopd.55*

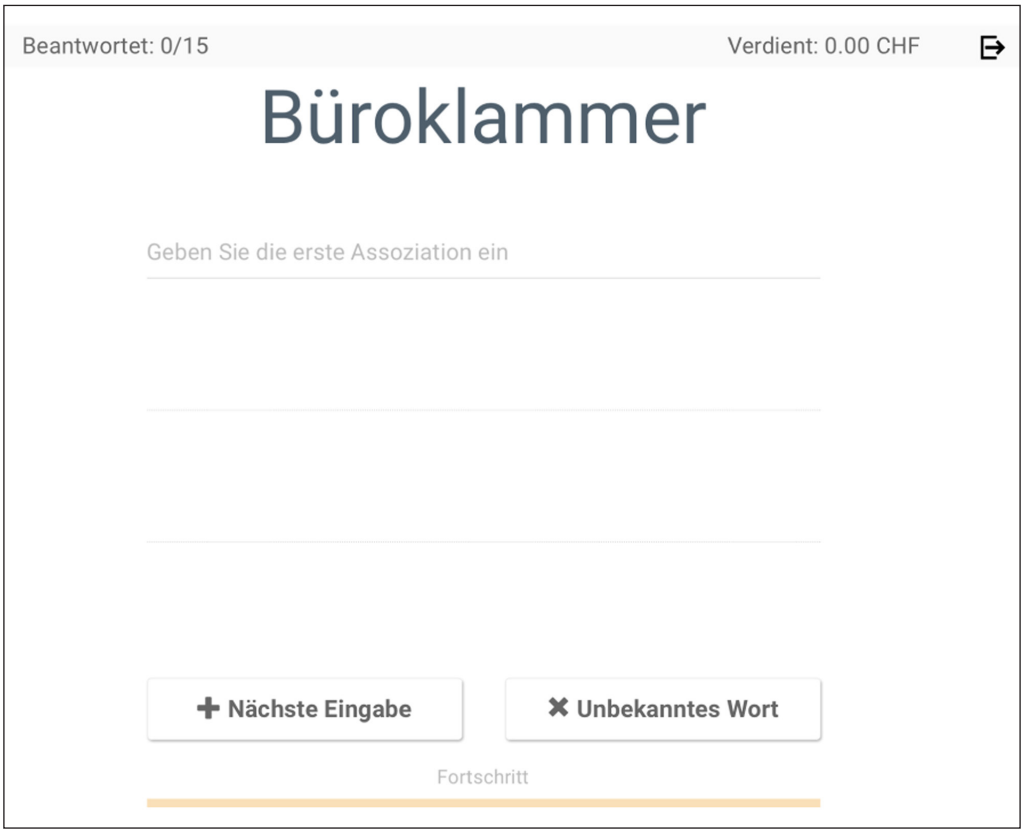

Figure 1 Screenshot of the free association task. The screenshot shows one trial requiring associations to the cue "Büroklammer" (paper clip) in the training mini study.

YOUNGER ADULTS				OLDER ADULTS			
24 YEARS	**27 YEARS**	**27 YEARS**	**28 YEARS**	**68 YEARS**	**68 YEARS**	**69 YEARS**	**70 YEARS**
man	**music**	**money**	**money**	human	Germany	**music**	state
music	**food**	write	**music**	country	instrument[1]	development	country
work	**work**	beautiful	**water**	animal	military	**money**	goods
change	**school**	**food**	Germany	**water**	**car**	exact	first name
water	**money**	**school**	**work**	occupation	**water**	wood	computer
family	clothing	read	**food**	**work**	country	make	instrument[1]
learn	write	important	furniture	**food**	**food**	**school**	family
wood	instrument[1]	economy	instrument[1]	**car**	name	computer	university
church	name	large	**love**	instrument	computer	children	church
love	summer	family	clothes	male	child	**work**	month

Table 1 Most Frequent Words in Association Task.

Note: Words were translated from German. Words in bold face are also among the ten most central words in the German (*money*, *music*, *work*, *school*, *food*, *water*, *car*, *love*, *green*, *important*) or English (*money*, *food*, *water*, *car*, *music*, *green*, *red*, *love*, *work*, *old*) SWOW data sets. [1] Musical instrument.

Cognitive assessment

The cognitive battery consisted of two sets of tasks fulfilling different purposes. The purpose of the first set was the assessment of people's general cognitive abilities and functioning. This set included a 20-minute timed version of the Advanced Progressive Matrices (APM; Hamel & Schmittmann, 2006) as a measure of general intelligence, a Digit-Symbol Substitution Test, as is found in the Wechsler Adult Intelligence Scale IV as subtest "coding" (WAIS-IV; Wechsler, 2008) as a measure of processing speed, the Mehrfachwahl-Wortschatz-Intelligenztest: Form I (MWT-A; Lehrl et al. 1995) as a measure of vocabulary size, and, finally, the DemTect (Kalbe et al., 2004) as a screen for dementia.

The purpose of the second set was to establish word-level links between the free association network and cognitive performance. This set included 10-minute category (animals) and phonemic (letter S) fluency tasks (e.g., Wulff et al., 2018, October 29), an episodic list memory task modeled after Penn Electrophysiology of Encoding and Retrieval Study (e.g., Healey & Kahana, 2016), and an associative recall task modeled after Naveh-Benjamin et al. (2003). Behavior in the two fluency tasks can be related to the free association network based on the fact that both cues and responses naturally included animals and words starting with the letters S. Participants retrieved between 62 and 113 animals and between 45 and 138 words starting with the letter S. The retrieved animals overlapped with 1.5% of cues and 0.8% of responses, whereas the retrieved words starting with the letter S overlapped with 11.1% of cues and 11.9% of responses. The episodic memory task and the associative recall task were populated with nouns from the cue set to establish comparability with the associative network. In the episodic memory task, a total of 20 lists of 16 words each were studied and subsequently recalled. Participants correctly recalled between 28.7% and 60.9% of words, with an additional 1.3% to 25% intrusions. In the associative recall task, 4 lists consisting of 16 word-pairs were presented and tested. Participants correctly recalled between 32.8% and 96.8% of pairs. See also ***Table 2*** for an overview of tasks included in the cognitive assessment in the MySWOW proof-of-concept study.

Entry and debriefing questionnaires

At study entry, participants provided demographic information concerning their primary language (German or Swiss German), their current occupation their highest academic degree, and the income level of their household. Participants further answered questions on their usual reading behavior, e.g., the number of books read in a year. At debriefing, participants were asked to provide information on their observations during the study, for example, whether they were able to sustain concentration while working on the free associations. The specific questions are reported in the codebook (see ***Table 3***).

PROCEDURE

Participants passing the initial screening over the phone were invited to to our laboratory at the University of

FILE	DESCRIPTION
participants.csv	Contains data on demographic information, reading behavior, debriefings survey, and all but four cognitive assessments.
associations.csv	Contains the corrected and uncorrected free association data.
episodic_memory.csv	Contains the episodic memory training and test data.
associative_recall.csv	Contains the associative recall training and test data.
animal_fluency.csv	Contains animal fluency response sequences.
letter_fluency.csv	Contains letter fluency response sequences.
codebook.pdf	Contains descriptions of all variable names in the data files.

Table 3 Description of Data Files.

TASK	DESCRIPTION	MOTIVATION	REFERENCE
Category fluency	Name all the animals you can in 10 minutes.	Predict performance from network	Wulff et al. (2018, October 29)
Phonemic fluency	Name all words starting with letter S you can in 10 minutes.	Predict performance from network	Griffiths et al. (2007)
Episodic memory task	Study a word list and then recall the words in any order (20 lists, 16 words per list).	Predict performance from network	Healey and Kahana (2016)
Associative recall task	Study a list of word pairs, then recall for each one word of a pair while being cued with the other (4 lists, of 16 word pairs).	Predict performance from network	Naveh-Benjamin et al. (2003)
Advanced Progressive Matrices	Solve abstract reasoning problems.	General cognitive abilities	Hamel and Schmittmann (2006)
Digit-Symbol Substitution	Assign digits to symbols according to rule.	General cognitive abilities	Wechsler (2008)
Mehrfachwahl-Wortschatz-Intelligenztest	Recognize words in list of words and non-words.	General cognitive abilities	Lehrl et al. (1995)
DemTect	Various cognitive tasks.	Screen for age-related pathologies	Kalbe et al. (2004)

Table 2 Tasks in the Cognitive Battery.

Basel for an introductory session lasting approximately 30 minutes. During this session, participants provided informed consent, completed the entry questionnaire, and were introduced to the web-based platform using a training mini-study involving 15 cues. Over the course of the next weeks, participants were instructed to log in and work on the free association task twice a day for 30 minutes each. On average, participants completed the free association task in 26.1 hours spread over 39.4 days. After completing the free association task, participants were invited back to the laboratory for a three-hour session that included the cognitive assessments and study debriefing.

The cognitive assessment and study debriefing session consisted of the following elements: First, participants filled out the debriefing questionnaire. Next, the verbal fluency tasks were conducted orally and recorded for later transcription by two student assistants responsible for data collection. Following the verbal fluency tasks, the participants were administered a 90-second timed Digit-Symbol Substitution Test in paper and pencil format. To conclude the first part of the lab session, the associative recall task was completed as a computerized task implemented in E-Prime (Psychology Software Tools, Inc., 2016) at a lab-computer. After a 10-minute break, the second part of the lab session began with the List Memory task, which was also implemented as a computerized task using E-Prime (Psychology Software Tools, Inc., 2016). The Mehrfachwahl-Wortschatz-Intelligenztest (MWT-A) was then conducted in paper and pencil format followed by a 20-minute timed version of the APM in paper and pencil format. The lab session concluded with the interactive verbal administration of the DemTect, carried out by one of the student assistants. Subsequently, participants received their monetary compensation for participation.

DATASET DESCRIPTION

Table 3 provides an overview of the different files containing the data. All data are available as comma-separated files. A *codebook.pdf* file provides descriptions of all variable names across the data files. All variable names and data labels have been translated to English. The association and fluency data, however, were not translated and are reported in German.

The data were published on the Open Science Framework (*10.17605/OSF.IO/VKWPS*) on 15 February 2021. The data are licensed under Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0) and follow the FAIR guidelines (Wilkinson et al., 2016).

QUALITY CONTROL

Free associations were obtained using a computerized task adapted from the German version of the large and long-running citizen-science project Small World of Words (De Deyne et al., 2019). Responses were cleaned in the following way. First, all responses matching either individual words or composites of words included in the German *aspell* dictionary were accepted as valid. The remaining words were subjected to manual correction. Overall, 4.2% of responses were corrected manually with a median string edit distance (i.e., the number of letters that were changed) of 2 (mean = 2.42). The data include the repeated sampling of a subset of words as an indicator to assess the reliability of free association. In the cognitive assessments, standardized and computerized tasks were used to improve comparability with previous work. All participants in the data set completed all tasks. Three additional participants started the study but dropped out (see Participants section).

ETHICAL CONSIDERATIONS

All data were recorded in reference to a random six letter identifier assigned to participants at the beginning of the study. Identifying information such as names or addresses was not recorded. Potentially identifying information such as participants' age, birthday, and occupation were not included in the publicly available files. Participants provided informed consent that included permission for public sharing of the data. The study was approved by the internal review board of the Department of Psychology at the University of Basel (# 014-17-1).

REUSE POTENTIAL

The reported data present the only publicly available resource containing large-scale free-association and cognitive performance data on the level of the individual (cf. Morais et al., 2013). The data can be reused in at least four different ways.

First, the data can be used to investigate individual and age differences in the structure of semantic networks. Past research has either used methods that prevented an assessment of large-scale networks or has analyzed semantic networks in the aggregate (e.g., similarity ratings, Wulff et al., 2018, October 29). There is only one other study that has elicited individual level semantic networks of comparable size (cf. Morais et al., 2013); however, it relied on a snowball approach that resulted in less comparable networks and its data are not publicly available. We present a first comparison of the structure of individuals' semantic networks in the companion paper (Wulff et al., 2022); however, that comparison considers only four structural properties and one approach to constructing the network from individuals' free associations. The current data could be used to explore other structural properties, such as