

**IDLaS-NL – A platform for running customized studies on individual differences in
Dutch language skills via the internet**

Florian Hintz¹, Olha Shkaravska¹, Marjolijn Dijkhuis¹, Vera van 't Hoff¹, Milou Huijsmans¹,
Robert C.A. van Dongen¹, Levi A.B. Voeteé¹, Paul Trilsbeek¹, James M. McQueen^{1,2},
& Antje S. Meyer^{1,2}

¹Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

²Radboud University, Nijmegen, The Netherlands

Running title: Customizable language test battery

Keywords: web-based testing; individual differences; language skills

Correspondence should be addressed to:

Florian Hintz

Max Planck Institute for Psycholinguistics

P.O. Box 310

6500 AH Nijmegen

The Netherlands

florian.hintz@mpi.nl

Word count abstract: max. 182 words

Word count body: 6279

Abstract

We introduce the Individual Differences in Language Skills (IDLaS-NL) web platform, which enables users to run studies on individual differences in Dutch language skills via the internet. IDLaS-NL consists of 35 behavioral tests, previously validated in participants aged between 18 and 30 years. The platform provides an intuitive graphical interface for users to select the tests they wish to include in their research, to divide these tests into different sessions and to determine their order. Moreover, for standardized administration the platform provides an application (an emulated browser) wherein the tests are run. Results can be retrieved by mouse click in the graphical interface and are provided as CSV-file output via email. Similarly, the graphical interface enables researchers to modify and delete their study configurations. IDLaS-NL is intended for researchers, clinicians, educators and in general anyone conducting fundamental research into language and general cognitive skills; it is not intended for diagnostic purposes. All platform services are free of charge. Here, we provide a description of its workings as well as instructions for using the platform. The IDLaS-NL platform can be accessed at www.mpi.nl/idlas-nl.

Introduction

Over the past decade, psycholinguistics has seen a growing interest in research on individual differences. Researchers have begun to acknowledge that comprehensive models must accommodate variability between language users, rather than focusing entirely on the average or group behavior. Indeed, according to theoretical views, individual-differences studies provide a powerful source of evidence bearing on key issues in the language sciences, such as the architecture of the language system and the mechanisms supporting language use (Kidd et al., 2018; Siegelman et al., 2017). As a result of this shift in thinking, the number of studies using individual-differences approaches has been steadily increasing (e.g., Dabrowska, 2018; Engelhardt et al., 2017; Favier et al., 2021; Isbilen et al., 2022; James et al., 2018; Johns et al., 2018; Li et al., 2019; McMurray et al., 2010; Schmidtke et al., 2018).

In spite of this positive trend, there are (at least) three reasons that might hold back labs from running studies on individual differences in language skills. The first reason relates to individual differences being best assessed using a multitude of tests for measuring the same underlying psychological construct (Miyake et al., 2000). Just as the vast majority of behavioral tests, tests measuring language skills suffer from the so-called ‘task impurity’ problem. Task impurity refers to the fact that performance on any single behavioral test is likely influenced by a multitude of skills (e.g., a speeded lexical decision task involves word recognition/lexical access *and* a speeded motor response). Thus, using a single test to gauge a psychological construct is likely to conflate the skill of interest with other skills. By using multiple tests that tap into the same underlying psychological construct but vary in their surface structure and/or response variable, researchers can apply statistical techniques that partial out unwanted variance and extract variance that is shared across the tests and reflect the skill of interest.

Second, since using language inherently involves general cognitive skills (e.g., non-verbal processing speed, Hintz et al., 2020a; Huettig & Janse, 2016; working memory, Baddeley, 2012; non-verbal reasoning, cf. Deary et al., 2007), variability in language skills should be characterized in concert with variability in general cognitive skills. This necessitates the inclusion of tests measuring the respective general cognitive skills involved in the language task(s) of interest.

The third reason is that individual-differences studies require large numbers of participants to achieve sufficient statistical power. As described by Schönbrodt and Perugini (2013), simple correlation coefficients stabilize at a sample size of 161 participants. As pointed out by Brysbaert (2019), for an effect size of $d = .4$ (corresponding to a correlation of $r = .2$, $p < .05$, two-tailed), 194 data pairs are required, which is much more than the number of participants typically tested for studies using factorial designs.

Taken together, large numbers of participants who each complete large numbers of tests amount to participant fees and testing time that many labs cannot afford. A critical bottleneck also concerns the man/woman power required for test administration and data pre-processing, in particular if the collected data involve manual transcription and annotation (but see Stark et al., 2022, for an alternative solution).

An alternative to lab-based test administration is remote testing via the internet, which is becoming increasingly popular. The availability of fast, flat-rate internet connections and affordable computer hardware for home use, as well as the host of open-source and commercial solutions for psychological testing have motivated many researchers to move their studies online. Indeed, as reported by Anwyl-Irvine et al. (2021), the number of papers tracked by Web of Science with the keywords ‘MTurk’ or ‘Mechanical Turk’ (Amazon’s platform for accessing online participants) increased from 121 publications in 2013 to 642 in 2018. Online testing speeds up data collection while yielding more diverse participant

samples than typically seen in lab-based testing (Garcia et al., 2022). Moreover, systematic comparisons of lab-based and online test administration (e.g., Cheung & Rensvold, 2002; Garcia et al., 2022; Germine et al., 2012; Ruiz et al., 2019; Ciecuch, 2015, for a tutorial) support the notion of measurement invariance. That is, while the absolute numbers may differ (e.g., reaction times recorded via the internet are likely longer than those recorded in the lab due to jitter in the internet connection and inferior hardware at home), the relationships between conditions within an experiment and the relationships between performance on different tests has been shown to be comparable (Hintz et al., in prep.). Given these advantages, online testing seems the perfect solution for alleviating some of the challenges properly powered and well-designed individual differences studies face.

Indeed, there are a number of existing commercial and open-access test batteries for assessing individual differences in language skills and skills related to language processing via the internet (e.g., Human Cognition Project, Morrison et al., 2015; PEBL, Mueller & Piper, 2014; Alberta Language Function Assessment Battery, Westbury 2006; ACS, Feenstra et al., 2018; PsyToolbox, Stoet, 2017)—some of which can be turned into customized batteries. However, to our knowledge, there is no solution yet that accommodates all three of the issues outlined above: (1) offering a test battery with multiple tests per psychological construct, (2) including tests measuring language *and* general cognitive skills involved in language processing, and (3) offering a comprehensive and user-friendly system for running these tests via the internet.

The present paper introduces the Individual Differences in Language Skills (IDLaS-NL) web platform that allows researchers to run customized studies on individual differences in language skills via the internet. The target language is Dutch and the target skill sets are word and sentence production and spoken word and sentence comprehension. We make available a set of 35 Dutch behavioral tests that were previously validated in participants aged

between 18 and 30 years of age. The tests are hosted on servers at the Max Planck Institute for Psycholinguistics. Selections of the tests can be combined into studies, which may consist of one or multiple sessions—depending on the researchers’ needs. Studies can be created, managed and deleted using an intuitive graphical user interface. We also make available a dedicated application (an emulated browser) wherein the studies are run to facilitate standardized test administration (i.e., each participant should use the same application). Upon start, the application takes up the full computer screen and thus reduces the likelihood of participants running other, potentially resource-consuming applications on the side. Each test has its own online database where the collected test data are stored. Results can be retrieved by mouse click within the graphical user interface. Item-level outputs, for some tests complemented with aggregated scores by participants, are made available via email. All services on the IDLaS-NL platform are free of charge.

In the remainder of this article, we provide more information about the tests and their validation. We introduce the elements of the web platform, including the IDLaS-NL website with useful information for researchers, the graphical user interface for creating, managing and deleting studies, and the test application. Finally, we describe how to use the platform and give some practical recommendations.

The Individual Differences in Language Skills Test Battery

Funded by the ‘Language in Interaction consortium’

(<https://www.languageininteraction.nl/>), IDLaS-NL was developed by researchers based at the Max Planck Institute for Psycholinguistics and at the Donders Institute for Brain, Cognition and Behavior between 2017 and 2022. To construct the battery, we conducted several pilot studies featuring parts of the tests (Brysbaert et al., 2021; Hintz et al., 2020a; Kapteijns & Hintz, 2021; Jongman et al., 2021) as well as one large pilot study that involved

testing 112 participants twice on all battery tests, with approximately one month's time in between, for assessing test-retest reliability (Hintz et al., 2020b; Hintz et al., 2022). Our working model (McQueen & Meyer, 2019) assumed that variability in linguistic processing skills (the skills at stake here are word- and sentence-level processing) is influenced by variability in linguistic experience (leading to linguistic knowledge) and general cognitive skills. Therefore, the tests in the battery measure the following psychological constructs: linguistic experience (with six tests), non-verbal processing speed (five tests), working memory (two tests), word production (four tests), sentence production (four tests), word comprehension (four tests), and sentence comprehension (four tests). To address the task-impurity problem, all psychological constructs are assessed with multiple tests. The final version of the test battery consists of 29 tests¹. Note that there are 6 more tests on the web platform (see below). Note also that in our main study (Hintz et al., in prep.), we also administered the Peabody Picture Vocabulary Test (Dunn & Dunn, 1997; Schlichting, 2005) and Raven's Advanced Progressive Matrices (Raven et al., 1998) as these tests are often regarded as the gold-standard for assessing receptive vocabulary size and non-verbal reasoning, respectively. However, since these tests are copyright protected, they are not made available (but see 'Additional tests' for open-access alternatives). Table 1 lists all tests, along with brief task descriptions. For more extensive descriptions, including materials, procedure, and descriptive statistics based on the norming sample of 748 participants, see Hintz et al. (in prep.).

¹ We count 'Gender cue identification' and 'Gender cue activation during sentence comprehension' as one test since in all of our studies we administered the identification test right before the sentence comprehension test. The main goals were to assess whether participants assign the expected gender to all pictures (which could later be used for screening participants) and to increase participants' sensitivity to gender information. Here, we provide them as separate tests in case other researchers do not wish the sentence comprehension test to be preceded by the identification test.

Additional tests

In addition to the tests that were part of our battery development efforts, we make available six tests that were previously developed and validated by researchers at Dutch, Belgian, English and US universities. The materials for these tests are free for use in scientific research and we used them for our own implementations of the test. The tests are listed under ‘Extra tests’ and ‘Extra tests production’, respectively, in the graphical user interface and measure the following psychological constructs: *Linguistic experience* (i.e., receptive vocabulary size, ‘Dutch Auditory & Image Vocabulary Test’, Bousard & Brysbaert, 2021; ‘Receptive vocabulary test (multiple choice)’, Vander Beken et al., 2018), *non-verbal reasoning* (‘Matrix reasoning test’, Chierchia et al., 2019), *personality traits* (‘BIG 5 personality traits’, Denissen et al., 2020), *word-reading* (‘One Minute Reading Test’, Callens et al., 2012), and *story-reading* (‘Story Reading Test’, Rouweler et al., 2020). We included these tests as alternatives for copyright-protected tests we previously used to assess receptive vocabulary size and non-verbal reasoning. The reading tests were included to capture the mediating influence that reading ability may have on spoken language processing (e.g., Huettig & Pickering, 2019). See Table 1 for brief descriptions of the additional tasks; for more information on the development and norming, see the paper(s) associated with each test.

Table 1: Overview of behavioral tests available on the IDLaS-NL platform.

Domain	Task	Duration	Task description
Linguistic experience	Stairs4Words (receptive vocabulary)	7 min	Adaptive vocabulary test based on a Yes/No decision task.
	Antonym production (productive vocabulary)	5 min	Participants hear a spoken word and are instructed to produce its antonym.
	Idiom recognition	3 min	Participants select the correct meaning for an idiomatic expression among four alternatives.
	Spelling test	5 min	Participants identify incorrectly spelled words in a list of 60 words (half of which are spelled incorrectly).
	Author recognition test	5 min	Participants identify fictional writers in a list of 132 names (90 are writers).
	Prescriptive grammar test	10 min	Participants carry out grammaticality judgements on spoken sentences featuring morpho-syntactic constructions known to be difficult for adult speakers of Dutch (e.g., ik vs. mij, als vs. dan, ze vs. hun).
Processing speed	Auditory simple reaction time test	3 min	Participants respond as quickly as possible to the presentation of a beep by pressing the space bar.
	Auditory choice reaction time test	4 min	Participants respond as quickly as possible to the presentation of one of two beeps (high or low) by pressing one of two buttons.
	Letter comparison test	5 min	Participants indicate as quickly as possible whether two letter strings are identical by pressing one of two buttons.
	Visual simple reaction time test	3 min	Participants respond as quickly as possible to the presentation of a geometrical shape by pressing the space bar.
	Visual choice reaction time test	4 min	Participants respond as quickly as possible to the presentation of one of two geometrical shapes by pressing one of two buttons.
Working memory	Digit span	7 min	Participants are instructed to recall sequences of digits in the order they were encountered (forward version) or in the reversed order (backward version).
	Corsi block clicking	7 min	Participants are instructed to recall sequences of identical spatially separated blocks by clicking on them in the order they were encountered (forward version) and in the reversed order (backward version).
Word production	Picture naming	7 min	Participants name pictures whose names vary in word frequency as quickly as possible.
	Rapid automatized naming	7 min	Participants are familiarized with four sets of five objects whose names vary orthogonally in frequency and neighborhood density. Each object set is arranged in an array consisting of five rows of six objects. Participants are instructed to name all objects in the array as quickly as possible.
	Verbal fluency	5 min	Participants name as many words as possible belonging to pre-specified categories (semantic part) and starting with letters provided ahead of time (phonological version) within one minute.
	Maximal speech rate	3 min	Participants are instructed to name the months of the year as quickly as possible.
Sentence production	Phrase generation	10 min	Participants are familiarized with a set of 16 common objects. These objects are presented in noun/adjectival phrases of increasing difficulty, which participants are instructed to name as quickly as possible.
	Sentence generation (structured)	12 min	Participants describe scenes depicting transitive actions (intransitives serve as fillers) as quickly as possible using pre-specified verbal material. Color-coding of the characters in the scenes coerces the production of active and passive sentences.
	Sentence generation (unstructured)	10 min	Participants describe scenes depicting transitive and intransitive actions and are free in their choice of words and syntactic structure/voice.
	Spontaneous speech	4 min	Participants speak freely for one minute about three topics provided ahead of time.
Word comprehension	Monitoring in noise in non-word lists	10 min	Participants monitor lists of non-words, presented in increasing levels of noise, for the occurrence of probe non-words, presented in the clear at the beginning of a trial.
	Rhyme judgement	5 min	Participants are presented with two non-words in succession and are instructed to judge as quickly as possible whether both non-words rhyme.
	Auditory lexical decision	7 min	Participants judge the lexicality status of an auditorily presented target word as quickly as possible.
	Semantic categorization	5 min	Participants judge as quickly as possible whether an auditorily presented target word belongs to a pre-specified semantic category.

Sentence comprehension	Monitoring in noise in sentences	10 min	Participants monitor predictable and non-predictable sentences, presented in increasing levels of noise, for the occurrence of probe words, presented in the clear at the beginning of a trial.
	Verb semantics activation during sentence comprehension	7 min	Participants are presented with two objects on the computer screen and a spoken sentence containing a target noun, which refers to one of the two objects. In half of the sentences, the target is predictable based on verb semantics. Participants indicate by button press which of the two objects is referred to in the sentence. On predictable trials, participants may respond before target noun onset.
	Gender cue identification	10 min	Participants indicate for 80 objects whether their names are de- or het-nouns (both common gender in English). <i>It is advised to run this test preceding the test 'Gender cue activation during sentence comprehension'.</i>
	Gender cue activation during sentence comprehension	10 min	Participants are presented with two objects (the same as in the Gender cue identification test) on the computer screen and a spoken sentence containing a target noun, which refers to one of the two objects. In half of the sentences, the target is predictable based on a determiner expressing the grammatical gender of the target noun. Participants indicate by button press which of the two objects is referred to in the sentence. On predictable trials, participants may respond before target noun onset. <i>It is advised to run this test following the test 'Gender cue identification'.</i>
	Self-paced reading	5 min	Participants read sentences of varying syntactic complexity in a self-paced fashion.
Extra tests	Dutch auditory & image vocabulary test	12 min	Participants hear a spoken word and select the picture associated with its meaning among four alternatives.
	Matrix reasoning test	8 min	Participants indicate which of four possible shapes completes a matrix of geometric patterns.
	BIG 5 personality traits	5 min	Participants rate personality statements.
	Receptive vocabulary test (multiple choice)	7 min	Participants read target words (varying in difficulty) and select the correct meaning among four written alternatives.
Extra tests production	One-minute reading test	3 min	Participants read as many words (increasing in difficulty) as possible within one minute.
	Story reading test	5 min	Participants read as much as possible of a story within three minutes.

Technical details of the IDLaS-NL web platform

Website

The website www.mpi.nl/idlas-nl provides a brief introduction to and overview of the IDLaS-NL platform. It describes the platform's most important features, provides some scientific background, a user manual, and an FAQ page. The website links through (link to be found in the 'User manual' section) to the graphical user interface where studies can be created, edited and deleted, and where results can be retrieved. The website's default language is English, but Dutch is available too.

Graphical user interface

The IDLaS-NL graphical user interface is implemented in PHP. The landing page provides a brief overview of the platform's functionality. To be able to access the subsequent page, users must provide an email address, accept the 'Terms of Use' and type in the name of the Dutch city where the Max Planck Institute for Psycholinguistics is located. This CAPTCHA is necessary to prevent bots from automatically creating large numbers of studies. When users select the tests they wish to include in their research, determine the number of test sessions and the order of tests within the session(s), their configurations are stored in a database. This study configuration is associated with the provided email address and a generated researcher key. Moreover, a unique study key is generated for the configuration and stored alongside the other information (see below for a more extensive description).

Framework for interactive experiments (Frinex)

All online tests were programmed in Frinex (Withers, 2016). Frinex has been designed and developed by Peter Withers and has been used at the Max Planck Institute for Psycholinguistics for online web experiments since 2015 and for offline field experiments since 2016. It is under active

development, which allows for custom features to be added for novel experiment requirements. Stimuli can be shown in Frinex in various forms such as text, audio or video. Visual stimuli can be animated so that the presentation includes movement on screen. Participant responses can be recorded in a number of ways, for example, by simple button clicks, rating buttons, textual input, recorded audio or video. Timing data is collected during the response period, which includes the time between key presses when textual input is used. All visual elements in the experiment including stimuli presentation can be customized with CSS (cascading style sheets) or with various predefined styles. Each of the 35 tests on the IDLaS-NL platform is a stand-alone experiment, with its own URL, and associated with a database that stores the data for that experiment only. The data are tagged with study and researcher keys, which enables the retrieval of specific portions of test data from an experiment database. When selecting multiple tests for a study, individual experiments are ‘chained together’ by listing them as ‘steps’ in the URL of the first experiment in the chain. This URL also contains the study and researcher keys, which enables the transmission of both keys from one experiment to the next and the use of both keys for tagging the collected data in each experiment.

Participant testing environment: The Electron application

In principle, all IDLaS-NL studies and individual experiments can be run in any browser. However, given the ever-changing nature of browsers to accommodate the latest technical advancements, for a standardized test administration we strongly recommend IDLaS-NL users to provide their participants with the dedicated application we make available.

This application is built using the Electron software framework (<https://www.electronjs.org/>), which is designed to create cross-platform desktop applications using web technology. It includes the Chromium browser engine that is also used in Google Chrome. The application can be run immediately after downloading (i.e., no separate installation procedure is

required) and can simply be moved to the trash once it is no longer needed. Windows and Mac versions are available².

Asking all participants to complete the study in the Electron application has several advantages: First, it ensures that all use the same (version of the) browser. Since we developed all tests for use in the Electron application, all experiment features should therefore work without problems. Second, we customized the application such that, upon start, it takes up the full screen size (address line, tabs, and irrelevant visual features have been removed), which reduces the likelihood of participants running other applications alongside the experiments that may consume the laptop's processing resources and/or internet bandwidth. Moreover, since the application takes up the full screen, participants are more likely to focus their attention on the tests since switching back and forth between distracting browser tabs is not possible. Finally, although the timing precision of trial events is subject to a number of influences (e.g., operating system, CPU, internet bandwidth, Anwyl-Irvine et al., 2021; Monen et al., in prep.), using the same browser in all participants eliminates one potential additional source of between-participant variability.

If unsolvable issues prevent running the Electron application on a participant's computer, the tests may also be carried out in a 'regular' browser (preferably Chrome). Instructions for such cases are provided in the PDF that is sent after submitting a study configuration. Throughout all test sessions, participants must maintain an active internet connection.

Hardware

The majority of tests require responses using the mouse or the keyboard. Speaking and listening tests on the IDLaS-NL platform require participants to use a microphone and headphones (integrated speakers work as well, but are not recommended). To ensure that the microphone works

² Just as other browsers, the Electron application leaves traces on participants' computers. These traces include timestamps, keys, hashes, and participant responses, and are neither harmful nor considered sensitive information. The FAQ page (<https://www.mpi.nl/idlas-nl/faq>) provides a simple tutorial for removing the traces after study completion.

properly, a test in the form of an additional Frinex experiment is automatically added to the beginning of each session that contains a test that requires speaking. These tests are tagged internally and are recognized automatically by the system and no action from the user is required. Before using the microphone in the Electron application, participants must give permission by clicking on the appropriate button in a pop-up dialog. The subsequent microphone test consists of participants naming four written Dutch words in a fixed order. Per word, a recording is made. Next, participants are presented with the just-made recordings, one at a time, in random order and have to click on the word they just heard themselves produce. After recognizing all four words successfully, participants can continue and start the test session; in case less than four words were recognized correctly, the process re-starts and is repeated until successful completion. Should the problem persist, participants need to contact their experimenter. Some potential solutions for solving microphone problems are listed on the FAQ page: <https://www.mpi.nl/idlas-nl/faq>.

Data retrieval

As mentioned above, all experimental data are tagged with study and researcher keys for later retrieval of the results. Results can be requested via the graphical user interface for each test individually (log on by providing email address and completing CAPTCHA, provide study and researcher keys and click on ‘Open existing study’). Once per hour, the system checks for which tests results have been requested and runs an R script to retrieve and process the requested data for those tests.

For language production tests, all recordings (.wav files) that match study and researcher keys are selected, zipped and made available for download. For all other tests, the data points that match study and researcher keys are downloaded from the Frinex server by the R script and are subsequently pre-processed to yield item-level CSV outputs (i.e., one line per participant per item). For some accuracy-based tests, the item-level output is complemented with aggregated summary

scores per participant. All resulting CSV output files are zipped, along with a PDF file listing the stimuli for that experiment and a PDF file providing a legend for the output files column headers, and made available for download. Additionally, the zip files containing recordings or CSV files always contain an irregularities CSV file, which lists information about potential problems that occurred during the test and/or data retrieval.

The R script eventually sends out an email (one per test) to the provided email address, containing a link that leads to the zip file of the respective test. Once produced, the download link and the associated zip file are available for seven days.

A step by step guide on how to use IDLaS-NL

The flowchart in Figure 1 illustrates the main steps for creating and running studies using the IDLaS-NL platform. Below, we provide more details on each of the steps and give practical recommendations based on our own experience, comprising more than 1000 online participants.

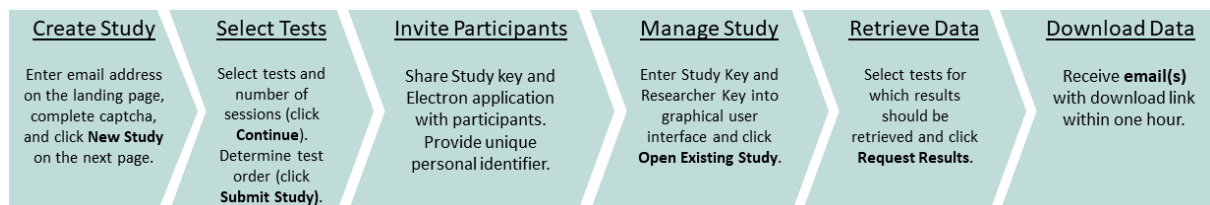


Figure 1: Main steps for creating and running studies using the IDLaS-NL platform.

Registration

The first step to using the IDLaS-NL platform is to register on the landing page. Be sure to use an email address with appropriate security levels. The email address functions as a ‘user name’ and anyone with access to the inbox could get access to the researcher key, which functions as a ‘password’. Users might want to add ‘IDLaS-NL team’ as a secure sender to prevent the email provider from regarding our emails as spam.

To advance to the study creation/modification sections of the platform, users must accept the Terms of Use (which includes the Data Processing Agreement; both in the Appendix).

Creating a study

By clicking on the ‘New study’ button, users initiate the creation of a new study. On the subsequent page, they can select the tests they wish to include in their study. The tests are grouped by the construct they measure. The tests that were not developed and piloted by the IDLaS-NL team, are highlighted as ‘extra tests’. The approximate duration of each test is provided in brackets; clicking on the “i” icon behind each test opens a pop-up featuring a short task description. At the bottom of the page, users can indicate the number of sessions into which they want to divide the tests. A test can only occur once in a session.

By clicking the ‘Continue’ button, users advance to the page where they can determine the order of tests within their session(s). They can change the number of tests per session by clicking the ‘- (remove)’ button and by clicking the ‘+ (add another test)’ button. The drop-down list next to each test number in a session features all tests selected on the previous page. Users determine the order of tests within a session by entering the tests into positions 1 through n. When satisfied with their configuration, users advance by clicking the ‘Continue’ button. The system then performs a check on the configuration. The following aspects are evaluated:

1. The same test occurs multiple times in one session: Users receive an *error* message and cannot create the study until the error is resolved. They need to adjust their configuration and click on the ‘Continue’ button at the bottom to trigger a new evaluation.
2. The same test occurs in different sessions: Users receive a *warning* message that can be ignored in case this is intentional (click on the ‘Continue’ button).
3. One or more previously selected tests do not occur in any of the sessions: Users receive a *warning* message that can be ignored if this is intentional (click on the ‘Continue’ button).

The study configuration is finalized by clicking the ‘Submit study’ button. Upon clicking that button, the study and researcher keys are generated and stored in the PHP database, along with the provided email address. In more technical terms, a URL for each session is created that contains the included tests as steps, study and the researcher keys and a placeholder for the participant ID. Shortly after submission, an automatic email is sent to the provided address, detailing the study configuration (sessions and tests) and listing study and researcher keys. The email also has a PDF file attachment, which contains, among others, the links where Windows and Mac versions of the Electron application can be downloaded.

Modifying and removing an existing study

If researchers wish to adjust an existing study configuration, they log on to the landing page by providing their email address and completing the CAPTCHA. Next, within the ‘Existing study’ section, they provide the study key for the study that should be modified and their researcher key and click on ‘Open existing study’. On the following page, users can adjust the study configuration. The graphical interface and configuration evaluation are the same as for creating a new study. The URL(s) as generated and stored when creating the study will be overwritten. Similarly, users may decide to delete their study configuration by clicking the corresponding button on the same page.

Note that study deletion refers to deleting the configuration as stored in the PHP database. It does not refer to the deletion of previously collected experimental data, tagged with the corresponding study key.

Running a study

In order for a participant to take part in a study, they need the following information: (1) the links for downloading the Windows/Mac version of the Electron application, (2) the study key, and (3) a personal identifier that is unique for that participant. This personal identifier can be assigned by the researcher or made up by the participant. The identifier should be at least three characters long and must not contain any personal information. In case of technical issues/questions, users and participants can consult the IDLaS-NL FAQ page: <https://www.mpi.nl/idlas-nl/faq>.

When inviting participants to take part in a study, users may want to inform them about the duration of the study, the number of sessions, the time that can/should be between two sessions, and the hardware required to complete the study such as headphones and/or a microphone. In case a study consists of multiple sessions, participants need to make sure to enter the same identifier at the beginning of each session.

When taking part in a study, participants open the Electron application. The first page prompts them for the study key and their personal identifier. When clicking on ‘Indienen’ (‘Enter’), the application retrieves the URL, or URLs in case of multiple sessions, associated with the entered study key. The subsequent page lists the sessions the participant still has to complete. Participants need to select the session they want to complete. The tests that are part of the selected session will then be run in succession. At the end of a session, participants are directed to the start screen where study key and personal identifier need to be entered. After entering this information, the remaining sessions are listed and participants select the session to be completed next.

It is important to highlight that closing the application during a test will mark that test as completed. If participants log on again at a later point in time, the application checks whether there is a test in a session that has not been completed yet and will offer to continue from there. For example, let us assume that a study consists of two sessions, which each consist of three tests. A participant closes the application in the middle of test 2 in session 1 and logs on the next day to complete the study by providing the study key and personal identifier. The application will list session 1 and session 2 as incomplete. When selecting session 1, the participant will be presented with test 3 from session 1. When selecting session 2, they will be presented with test 1 from session 2. It is therefore crucial to provide participants with precise information on the time intervals between tests and between sessions and to highlight that once started, a test must be completed in one sitting.

In most studies demographic information about the participants is collected. Given the precise research question, these questionnaires can drastically vary in length. Here, we decided to only collect the most basic information: At the beginning of the first session of each study, participants are asked to provide their age, gender (female, male, other), native language, handedness, and any medical issues (if applicable *and* if they want to share). When requesting the results for an experiment from the first session, a separate email is sent containing the results of the short questionnaire. Note that in terms of databases there is strict separation of anonymized experimental data and anonymized questionnaire data. Users are invited to complement this basic set of questions using a separate (online) survey.

Solutions to common technical problems

In spite of our best efforts to debug all components of the IDLaS-NL platform, things can go wrong. We strongly advise researchers to test the battery a number of times, e.g. by sending the link to colleagues using different types of computers and browsers. Our own experience has shown that

most errors occur (1) during data collection, for example when trying to complete the microphone test but the application has no permission to use the microphone or the application does have permission but no speech is recorded and the test cannot be completed. In such cases, it makes sense to test whether the microphone makes recordings at all using a different application. One may also check the operating system settings and manually give permission to the application to use the microphone. The other place where we sometimes see errors is (2) when retrieving results. In such cases, the email that would otherwise contain the link from where the zipped result file can be downloaded features an error message. To give one example, the error message might state that there are no data yet to be retrieved (i.e., no participants have yet finished the specific test). Other error messages may be more complex. When users receive an error message, they should try again at a later point in time. Should the error persist, they may send an email including the provided error message to the developers of IDLaS-NL (idas-nl@mpi.nl). Please note that we cannot support individual participants experiencing technical issues.

Practical recommendations

Throughout the article, we have already provided some recommendations, however, there are four topics that warrant to be singled out. The first recommendation is a mundane one but may have a substantial influence on the quality of the collected data: Users should be sure to instruct their participants to complete the studies in a quiet environment and on their own. A quiet surrounding will massively improve the quality of speech recordings and will, of course, increase the performance of the participant as compared to a noisy environment.

Second if at all possible participants should complete a study with multiple sessions on the same computer. While the system is capable of handling the same participant completing different sessions on different devices, we do not recommend that—certainly not if tests are included whose dependent variable is time-critical (e.g., auditory lexical decision). The reason is that much of the

jitter we see in timing precision of trial events (auditory playback, response logging, Monen et al., in prep.) is associated with features of the hardware that was used (e.g., the keyboard's polling rate). Using the same computer across different sessions eliminates a potential source of noise, which is crucial for individual-differences studies.

Third, and most importantly, users must pay close attention to the ethical consent procedure. To be clear, we provide the IDLaS-NL web platform as a technical service, but it is the users' responsibility to ensure that they have sufficient coverage for collecting data using our services. As with other web platforms that can be used for running online experiments, the collected data are stored on our servers. Thus, users must have ethical coverage that approves of storing online experimental data from human participants on European (i.e., Dutch and German) servers. Furthermore, users must implement a way of collecting informed consent (in line with their ethical coverage) from their participants (e.g., a separate online survey, a signed PDF).

Further information

One of the challenges that experiments conducted via the internet face, compared to conducted in the lab, concerns the timing precision of trial events. In lab studies, the hardware is typically optimized for chronometric experimentation with jitters often below 10 ms across different stimulus delivery and experiment control environments (Bridges et al., 2020). The same timing precision cannot be achieved in online experiments since the experiments are delivered via the internet and variation in bandwidth influences timing precision. Similarly, the hardware the participants use is often inferior to that in the labs. Previous investigations into the jitter of delivering visual stimuli revealed imprecisions in stimulus duration ranging from -6.24 to 26.02 ms on average (across different platforms, Anwyl-Irvine et al., 2021). The same study found that response logging (RT delay calculated as the difference between known and recorded RT) ranged from 71.33 to 87.40 ms on average (across different platforms).

We conducted a similar study to assess the timing (im)precision in our IDLaS-NL test battery, as programmed in our own online environment (Frinex). For six battery tests that have a time-critical dependent variable, we assessed the jitter in presenting visual and auditory stimuli, recording a keyboard button press and recording human speech. We refer the reader to a description of that study, detailed in Monen et al. (in prep.). Overall the values are comparable to those reported in Anwyl-Irvine et al. (2021) and suggest that, while users have to accept some imprecision in timing, the overall pattern looks very promising.

Concluding comments

IDLAS-NL was developed for Dutch, and the linguistic tasks are consequently suitable only for testing speakers of that language. We are currently developing a German version, and in the near future plan to develop an English version and we would be delighted to hear from any researchers interested in being involved with us in developing these versions and/or versions for other languages.

Acknowledgements

This research was supported by the Netherlands Organization for Scientific Research (NWO), Gravitation grant ‘Language in Interaction’ (grant number 024.001.006).

Open Practices Statement

There are no data associated with the present manuscript.

References

- Anwyl-Irvine, A., Dalmaijer, E. S., Hodges, N., & Evershed, J. K. (2021). Realistic precision and accuracy of online experiment platforms, web browsers, and devices. *Behavior Research Methods*, *53*(4), 1407–1425.
- Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annu Rev Psychol*, *63*, 1–29.
- Bousard, I., & Brysbaert, M. (2021). The Dutch Auditory & Image Vocabulary Test (DAIVT): A New Dutch Receptive Vocabulary Test for Students. *Psychologica Belgica*, *61*(1), 1–17.
- Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. W. (2020). The timing mega-study: Comparing a range of experiment generators, both lab-based and online. *PeerJ*, *8*, e9414.
- Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition*, *2*(1), 16.
- Brysbaert, M., Sui, L., Dirix, N., & Hintz, F. (2020). Dutch Author Recognition Test. *Journal of Cognition*, *3*(1), 6.
- Callens, M., Tops, W., & Brysbaert, M. (2012). Cognitive profile of students who enter higher education with an indication of dyslexia. *PLOS ONE*, *7*(6), e38081.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233–255.
- Chierchia, G., Fuhrmann, D., Knoll, L. J., Pi-Sunyer, B. P., Sakhardande, A. L., & Blakemore, S.-J. (2019). The matrix reasoning item bank (MaRs-IB): Novel, open-access abstract reasoning items for adolescents and adults. *Royal Society Open Science*, *6*(10), 190232.
- Cieciuch, J., & Davidov, E. (2015). Establishing measurement invariance across online and offline samples. A tutorial with the software packages Amos and MPlus. *Studia Psychologica: Theoria et Praxis*, *2*(15), 83–99.

- Dąbrowska, E. (2018). Experience, aptitude and individual differences in native language ultimate attainment. *Cognition*, *178*, 222–235.
- Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*(1), 13–21.
- Denissen, J. J., Geenen, R., Soto, C. J., John, O. P., & Van Aken, M. A. (2020). The Big Five Inventory–2: Replication of psychometric properties in a Dutch adaptation and first evidence for the discriminant predictive validity of the facet scales. *Journal of Personality Assessment*, *102*(3), 309–324.
- Dunn, L. M., & Dunn, D. (1997). *Peabody Picture Vocabulary Test (3rd Edition)*. American Guidance Service.
- Engelhardt, P. E., Nigg, J. T., & Ferreira, F. (2017). Executive function and intelligence in the resolution of temporary syntactic ambiguity: An individual differences investigation. *The Quarterly Journal of Experimental Psychology*, *70*(7), 1263–1281.
- Favier, S., Meyer, A. S., & Huettig, F. (2021). Literacy can enhance syntactic prediction in spoken language processing. *Journal of Experimental Psychology: General*, *150*(10), 2167–2174.
- Feenstra, H. E. M., Murre, J. M. J., Vermeulen, I. E., Kieffer, J. M., & Schagen, S. B. (2018). Reliability and validity of a self-administered tool for online neuropsychological testing: The Amsterdam Cognition Scan. *Journal of Clinical and Experimental Neuropsychology*, *40*(3), 253–273.
- Garcia, R., Roeser, J., & Kidd, E. (2022). Online data collection to address language sampling bias: Lessons from the COVID-19 pandemic. *Linguistics Vanguard*.
<https://doi.org/10.1515/lingvan-2021-0040>
- Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin & Review*, *19*(5), 847–857.

- Hintz, F., Dijkhuis, M., van 't Hoff, V., McQueen, J. M., & Meyer, A. S. (2020b). A behavioural dataset for studying individual differences in language skills. *Scientific Data*, 7(1), 429. <https://doi.org/10.1038/s41597-020-00758-x>
- Hintz, F., Jongman, S. R., Dijkhuis, M., van 't Hoff, V., McQueen, J. M., & Meyer, A. S. (2020a). Shared lexical access processes in speaking and listening? An individual differences study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(6), 1048–1063.
- Hintz, F., Kievit, R. A., McQueen, J. M., & Meyer, A. S. (in prep.). *The principal dimensions of speaking and listening skills*.
- Hintz, F., Voeten, C., McQueen, J., & Meyer, A. (2022). Quantifying the relationships between linguistic experience, general cognitive skills and linguistic processing skills. In J. Culbertson, A. Perfors, H. Rabagliati, & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society (CogSci 2022)* (pp. 2491–2496). Cognitive Science Society.
- Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, 31(1), 80–93.
- Huettig, F., & Pickering, M. J. (2019). Literacy Advantages Beyond Reading: Prediction of Spoken Language. *Trends in Cognitive Sciences*, 23(6), 464–475.
- Isbilen, E. S., McCauley, S. M., & Christiansen, M. H. (2022). Individual differences in artificial and natural language statistical learning. *Cognition*, 225, 105123.
- James, A. N., Fraundorf, S. H., Lee, E.-K., & Watson, D. G. (2018). Individual differences in syntactic processing: Is there evidence for reader-text interactions? *Journal of Memory and Language*, 102, 155–181.
- Johns, C. L., Jahn, A. A., Jones, H. R., Kush, D., Molfese, P. J., Van Dyke, J. A., Magnuson, J. S., Tabor, W., Mencl, W. E., & Shankweiler, D. P. (2018). Individual differences in decoding

- skill, print exposure, and cortical structure in young adults. *Language, Cognition and Neuroscience*, 33(10), 1275–1295.
- Jongman, S. R., Khoe, Y. H., & Hintz, F. (2021). Vocabulary Size Influences Spontaneous Speech in Native Language Users: Validating the Use of Automatic Speech Recognition in Individual Differences Research. *Language and Speech*, 64(1), 35–51.
- Kaptein, B., & Hintz, F. (2021). Comparing predictors of sentence self-paced reading times: Syntactic complexity versus transitional probability metrics. *PLOS ONE*, 16(7), e0254546.
- Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual Differences in Language Acquisition and Processing. *Trends in Cognitive Sciences*, 22(2), 154–169.
- Li, M. Y. C., Braze, D., Kukona, A., Johns, C. L., Tabor, W., Van Dyke, J. A., Mencl, W. E., Shankweiler, D. P., Pugh, K. R., & Magnuson, J. S. (2019). Individual differences in subphonemic sensitivity and phonological skills. *Journal of Memory and Language*, 107, 195–215.
- McMurray, B., Samelson, V. M., Lee, S. H., & Tomblin, J. B. (2010). Individual differences in online spoken word recognition: Implications for SLI. *Cognitive Psychology*, 60(1), 1–39.
- McQueen, J. M., & Meyer, A. S. (2019). Key issues and future directions: Towards a comprehensive cognitive architecture for language use. In P. Hagoort (Ed.), *Human language: From genes and brain to behavior* (pp. 85–96). MIT Press.
- Miyake, A., Emerson, M. J., & Friedman, N. P. (2000). Assessment of executive functions in clinical setting: Problems and recommendations. *Seminars in Speech and Language*, 21(2), 169–183.
- Monen, J., Shkaravska, O., Withers, P., van den Heuvel, M., Weustiniak, J., Dirksmeyer, R., & Hintz, F. (in prep.). *Timing precision for online studies using the Individual Differences in Language Skills (IDLAS-NL) test battery.*

- Morrison, G. E., Simone, C. M., Ng, N. F., & Hardy, J. L. (2015). Reliability and validity of the NeuroCognitive Performance Test, a web-based neuropsychological assessment. *Frontiers in Psychology*, 1652.
- Mueller, S. T., & Piper, B. J. (2014). The psychology experiment building language (PEBL) and PEBL test battery. *Journal of Neuroscience Methods*, 222, 250–259.
- Raven, J., Raven, J. C., & Court, J. H. (1998). *Raven manual section 4: Advanced progressive matrices*. Oxford, UK: Oxford Psychologists Press.
- Rouweler, L., Varkevisser, N., Brysbaert, M., Maassen, B., & Tops, W. (2020). The Flamingo test: A new diagnostic instrument for dyslexia in Dutch higher education students. *European Journal of Special Needs Education*, 35(4), 529–543.
- Ruiz, S., Chen, X., Rebuschat, P., & Meurers, D. (2019). Measuring individual differences in cognitive abilities in the lab and on the web. *PLOS ONE*, 14(12), e0226217.
- Schlichting, L. (2005). *Peabody Picture Vocabulary Test Dutch-III-NL*. Amsterdam, NL: Harcourt Assessment BV.
- Schmidtke, D., Van Dyke, J. A., & Kuperman, V. (2018). Individual variability in the semantic processing of English compound words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(3), 421–439.
- Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, 47(5), 609–612.
- Siegelman, N., Bogaerts, L., Christiansen, M. H., & Frost, R. (2017). Towards a theory of individual differences in statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1711), 20160059.
- Stark, K., van Scherpenberg, C., Obrig, H., & Abdel Rahman, R. (2022). Web-based language production experiments: Semantic interference assessment is robust for spoken and typed response modalities. *Behavior Research Methods*, 1–27.

- Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology, 44*(1), 24–31.
- Vander Beken, H., & Brysbaert, M. (2018). Studying texts in a second language: The importance of test type. *Bilingualism: Language and Cognition, 21*(5), 1062–1074.
- Withers, P. (2016). *Frinex: Framework for Interactive Experiments*. Poster presented at the IT-Verantwortlichen-Treffen, Berlin, Germany.
- Westbury, C. (2006). The Alberta language function assessment battery. *Brain and Language, 99*(1), 53–54.

November 14, 2022

Individual Differences in Language Skills (IDLaS-NL) web platform

Terms of Use, V 1.0 (10/2020)

By using the online implementation of the Individual Differences in Language Skills Test Battery (hereafter "IDLaS"), you acknowledge that you have read these terms and conditions, understand them, and agree to be bound by them. If you do not agree with these terms and conditions, you must not use IDLaS. Please read the following Terms of Use and any accompanying documentation, carefully before you use IDLaS.

1. Contents and availability

The use of IDLaS is licensed to You by the Max Planck Institute for Psycholinguistics, with offices at Wundtlaan 1, 6525 XD Nijmegen, The Netherlands (Licensor), subject to these Terms of Use.

The Licensor wishes to allow You access to IDLaS, an online platform for conducting behavioural studies on individual differences in language skills, developed by the Psychology of Language Department at Max Planck Institute for Psycholinguistics, funded by a Gravitation grant from the Dutch Government (NWO).

The contents and functionality of IDLaS have been examined carefully. The Licensor has made reasonable efforts in designing, testing and implementing the behavioural tests made available, to the best of their knowledge. Nonetheless, the Licensor does not warrant the use of IDLaS to be complete, up-to-date or constantly available. The Licensor does not bear any responsibility for third-party contents, for the use of IDLaS different to the original one (see purpose below), for the accessibility via external links, nor for their availability. Furthermore, the Licensor waives any responsibility for contents that may be illegal or violate common decency.

2. Purpose

The Licensor has built IDLaS for the purpose of providing interested researchers, clinicians and professionals in education with a tool to assess individual differences in language skills. IDLaS encompasses a collection of 36 behavioural tests, which may be used in different combinations and orders via the internet. IDLaS must not be used for diagnosing any kind of (developmental) impairments or disorder. You hereby agree to use IDLaS and data obtained through it exclusively for the purpose of scientific research or education under an academic, research, government, health, or commercial entity (but in the last case, solely for non-commercial purposes).

3. Rules of use

You accept to use IDLaS non-commercially and for the aforementioned purpose. You are not authorized to make the contents of the use of IDLaS available to third parties. Furthermore, by accepting the present Terms of Use, You verify that You maintain an ethics committee or Institutional Review Board with European or equivalent standards that reviews and approves research involving human subjects and that the person accepting the Terms of Use has the authority to do so (i.e., accepting the responsibility for its actions related to the use of IDLaS). By accepting the Terms of Use, You agree that all studies conducted by You using IDLaS have been approved by your Institution's ethics committee or Institutional Review Board before beginning data collection and that experimenters have had ethics training addressing human subjects policy and issues. In particular, the review board must provide coverage for the storage of pseudonomized and partially non-anonymous (audio recordings) data on the Licensor's server. It is Your responsibility to collect informed consent from the test takers before starting data collection.

November 14, 2022

You accept that the Licensor may use the data collected in the realm of your study for the purpose of scientific research or education under an academic, research, government, health, or commercial entity (but in the last case, solely for non-commercial purposes).

4. Proprietary rights to the contents of the use of IDLaS

You agree to respect copyrights, rights to names, trademarks and other intellectual property rights, if any, of the Licensor and of third parties when making use of the online implementation of the test battery. By enabling access to the use of IDLaS, the Licensor does not grant any licence nor any other right of use.

5. Improper use of the use of the Application IDLaS

You agree to refrain from any improper use of the online implementation of the test battery; in particular, no security precautions must be circumvented. IDLaS may not be used to create fake, libelous, misleading, or defamatory content of any kind.

Furthermore, no facilities may be used nor may any applications be run that could lead to a damage or a performance failure of any of the facilities from which the use of IDLaS is provided, in particular through changes in the physical or logical structure of the servers or its network or of any other network. No commercial, systematic use of the use of IDLaS is permitted without the Licensor's consent.

6. Limitation of Liability (for Civil Law Countries)

Without prejudice to Licensor's responsibility for tort or for violation of mandatory laws of the Federal Republic of Germany, the Licensor shall not be liable for any damages You incur when using the online implementation of the test battery.

7. Representation, Warranty, Limitation of Liability, Indemnification (for Common Law Countries)

7.1 Representation and Warranty

LICENSOR REPRESENTS THAT LICENSOR HAS ALL RIGHTS REQUIRED TO MAKE AVAILABLE AND DISTRIBUTE THE MATERIALS. EXCEPT FOR SUCH REPRESENTATION, THE MATERIALS ARE PROVIDED "AS IS" AND "AS AVAILABLE" AND WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, NON-INFRINGEMENT, MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, AND ANY WARRANTIES IMPLIED BY ANY COURSE OF PERFORMANCE OR USAGE OF TRADE, ALL OF WHICH ARE EXPRESSLY DISCLAIMED.

WITHOUT LIMITING THE FOREGOING, LICENSOR DOES NOT WARRANT THAT: (A) THE MATERIALS ARE ACCURATE, COMPLETE, RELIABLE OR CORRECT; (B) THE MATERIALS FILES WILL BE SECURE; (C) THE MATERIALS WILL BE AVAILABLE AT ANY PARTICULAR TIME OR LOCATION; (D) ANY DEFECTS OR ERRORS WILL BE CORRECTED; (E) THE MATERIALS AND ACCOMPANYING FILES ARE FREE OF VIRUSES OR OTHER HARMFUL COMPONENTS; OR (F) THE RESULTS OF USING THE MATERIALS WILL MEET DOWNLOADER'S REQUIREMENTS. DOWNLOADER'S USE OF THE MATERIALS IS SOLELY AT DOWNLOADER'S OWN RISK.

7.2 Limitation of Liability

IN NO EVENT SHALL LICENSOR BE LIABLE UNDER CONTRACT, TORT, STRICT LIABILITY, NEGLIGENCE OR ANY OTHER LEGAL THEORY WITH RESPECT TO THE MATERIALS (I) FOR ANY DIRECT DAMAGES, OR (II) FOR ANY LOST PROFITS OR

November 14, 2022

SPECIAL, INDIRECT, INCIDENTAL, PUNITIVE, OR CONSEQUENTIAL DAMAGES OF ANY KIND WHATSOEVER.

7.3 Indemnification

You will indemnify and hold Licensor harmless from and against any and all loss, cost, expense, liability, or damage, including, without limitation, all reasonable attorneys' fees and court costs, arising from i) Your misuse of the use of IDLaS; (ii) Your violation of these Terms of Use; or (iii) infringement by You or any third party of any intellectual property or other right of any person or entity contained in the use of the online implementation of the test battery. Such losses, costs, expenses, damages, or liabilities shall include, without limitation, all actual, general, special, and consequential damages.

8. Data Protection

By using IDLaS You, the data controller, commission the Max Planck Institute for Psycholinguistics, with offices at Wundtlaan 1, 6525 XD Nijmegen, The Netherlands, commissioned processor, to process data on your behalf. Please see our **Data Processing Agreement** for the detailed provisions. By accepting the terms of use, you also accept the data processing agreement.

By using your data, the Max Planck Institute for Psycholinguistics is data controller and is abiding to data protection laws.



Commissioned data processing agreement in conformity with Art. 28 GDPR

by and between

Data controller/You/user of IDLaS (subsequently referred to as the controller)

and

Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V. (subsequently referred to as „MPG“)

Represented by the Managing Director Prof. Dr. Antje Meyer acting on behalf of the

Max Planck Institute for Psycholinguistics

Wundtlaan1
6525XD Nijmegen
The Netherlands

- subsequently referred to as „**MPI**“ -
- commissioned processor -
- subsequently referred to as the processor -

Section 1 Subject Matter and Duration of Commission

The processor makes available a collection of 36 behavioural tests suited to assess individual differences in language skills. The tests have been piloted in individuals aged between 18 and 30. The tests require participants to provide spoken and written responses, as well as (speeded) button presses. The processor provides an online platform for controllers to run the tests remotely (i.e., via the internet). Controllers may create customized versions of the test battery by including/excluding tests of their choice, in the order of preference and dividing tests into sessions of variable length. A session-test configuration is labelled 'a study', associated with a unique identifier ('study key') and stored. The Controllers can manage studies associated with the provided email address and identifier ('researcher key'). Managing includes revising and deleting a session-test configuration. The processor also makes available the 'Electron program', wherein the tests can be run. Electron is an emulation of the Chrome browser. Finally, the processor provides a 'data retrieval' service, which includes a set of R scripts that retrieve and aggregate the collected data associated with a provided study key.

In its initial phase, the IDLaS-NL service and technical support will be available until 31.05.2026.



Within the scope of using our service, the data collected and uploaded to our server will be stored and made available to the controller who used the service.

The processor shall perform the agreed services only as defined in the relevant agreements and in accordance with the subject matter and duration of the commission.

Section 2 Scope, Type and Purpose of the Intended Data Processing

- (1) Within the scope of the commission, the processor shall process the following personal data: *Email address, audio (voice) recordings, age, gender, educational background, mother tongue*
The personal data relate to the following groups of individuals:
 - *E-Mail address: Controller*
 - *Age, gender, educational background, mother tongue, medical issues: Any test taker of the IDLaS battery*
 - *Audio (voice) recordings: Any test taker carrying out a language production test of the IDLaS battery*
- (2) Data processing shall be performed for the following purposes only: *To make use of our service/tool. The processor reserves the right to use the collected data for academic purposes, including – but not limited to – statistical analyses and the publication of scientific articles.*
- (3) According to the controller's Security Levels Concept, the data are classified in the following protection categories:
 - normal
 - high**
 - very high
- (4) The processor may only collect, process or use the personal data within the scope of the documented instructions given by the controller. The processor shall be bound by the controller's instructions throughout the life of the contract.
- (5) Pursuant to this Agreement, all contractual data may only be processed and stored in countries which are members of the European Union or signatory states to the Agreement on the European Economic Area. The processor hereby gives their assurance that they will protect the contractual data against access by governmental entities outside the European Union or the European Economic Area.
- (6) Any transfer of personal data to a third country requires the controller's prior approval and shall only take place if the special conditions laid down in Art. 44 ff. General Data Protection Regu-



lation (GDPR) are complied with. These shall be specified separately.

- (7) The processor shall refrain from using data which are disclosed to them during or in the context of fulfilment of the contract for purposes other than those stipulated. Copies or duplicates must not be made unless the controller is aware of this and has given prior written permission. Exception: backup copies which are made so that data processing can be performed properly and so that liability and warranty claims can be fulfilled.
- (8) The processor may not provide information to third parties, including data protection supervisory authorities, without consulting the controller first.

Section 3 Controller's Rights and Duties

- (1) The controller is the data controller (§ 4, para. 7 GDPR) for commissioned data processing carried out by the processor. Assessment of the legality of the data processing is in the controller's responsibility.
- (2) The controller shall be responsible for upholding the rights of the parties involved. The processor must immediately notify the controller if parties involved assert their rights vis-a-vis the controller.
- (3) The controller shall instruct the processor to comply with all obligations under this Agreement. The controller shall be entitled to issue additional instructions to the processor at any time regarding the type, scope and nature of data processing. Instructions may be issued in writing or via email bearing a digital signature.

Section 4 Processor's Rights and Duties

- (1) The processor regularly controls data processing and internal processes and immediately informs the controller in case of suspected violation of the protection of personal data, cases of serious operational disruptions or other irregularities in the processing of the controller's data. The processor shall take the necessary measures to secure the data and to mitigate possible adverse consequences of the persons concerned and shall consult with the controller without delay. The processor shall immediately provide the controller with all the information requested, in particular the information required for the reporting in ac-



cordance with Articles 33, 34 GDPR and shall assist it in the fulfilment of the controller's obligations under Articles 33, 34 GDPR. The processor shall allow the controller or third parties designated by the controller without delay to carry out their own investigations of data processing.

- (2) The processor shall immediately inform the controller if the processor deems an instruction to be in violation of the GDPR or other data protection provisions of the EU or the member states.
- (3) The processor shall notify the controller before announced inspections by the data protection authorities if contractually agreed services will be affected or if the data protection authorities' inspection might have consequences for the type and manner of contractual fulfilment. The processor shall also notify the controller if an authority investigates the processing of personal data by the processor in the course of criminal proceedings or proceedings for fines or if investigations for other reasons are carried out in connection with such data on the part of the processor.
- (4) Furthermore, when first requested to do so, the processor shall immediately and comprehensively supply the controller with all necessary information regarding the collection, storage, processing or transfer of personal data which may be needed to fulfil any duties to provide information vis-a-vis parties involved or the relevant authorities. The processor shall assist the controller in the fulfilment of their obligations under Articles 35 and 36 GDPR (data protection impact assessment and prior consultation) and shall provide all information necessary for this purpose.
- (5) The processor shall support the controller to the best of their ability in proceedings before the supervisory authority, in fine, criminal or administrative proceedings, in disputes with affected parties or third parties in connection with commissioned processing or personal data, in particular in the event of a claim for possible claims pursuant to Article 82 GDPR. As far as these activities exceed the contractually agreed scope of services, the processor can demand an appropriate remuneration.
- (6) The support services set out in paragraph 1 to 5 are provided by the processor free of charge insofar as is reasonable and appropriate.
- (7) Correction, deletion or blocking of personal data shall only be performed by the processor upon appropriate instruction and/or prior approval from the controller. The processor must set up their infrastructure in such a way and take all other measures



necessary so that they can immediately fulfil such requirements issued by the controller. Within the scope of the processor's possibilities and within the scope of the instructions of the controller, the processor supports the controller in fulfilling the inquiries and claims of persons concerned in accordance with Chapter III of the GDPR.

- (8) The processor assures that they have appointed a competent and reliable data protection officer who is granted the necessary time to carry out their tasks in accordance with Articles 38 and 39 GDPR. The processor provides the controller with the data protection officer's contact details without delay. This applies also in the event that a new data protection officer takes office. If the processor is not obliged to appoint a data protection officer, they shall inform the controller accordingly.

§ 5 Commitment to Data Confidentiality

- (1) The processor guarantees that the employees involved in processing the data of the principal and other persons working for the processor are prohibited from processing the data outside the instructions of the controller. Furthermore, the processor guarantees that the persons authorized to process the personal data have undertaken to maintain confidentiality or are subject to an appropriate statutory commitment to non-disclosure. The commitment to confidentiality/non-disclosure shall continue to apply even after termination of the commission.
- (2) The processor hereby gives their assurance that they have made the staff assigned to the commission familiar with the data protection legislation relevant to them. The processor shall monitor compliance with data protection legislation and with the instructions given.

Section 6 Technical and Organizational Measures

- (1) The processor shall design the internal organization within their area of responsibility in such a way that it meets the special requirements of data protection. They will take technical and organizational measures to adequately protect the controller's data in order to ensure a protection level appropriate to the risk, based on the level of data protection specified by the client.
- (2) The technical and organizational measures must meet the requirements of the GDPR (Art. 32 GDPR) and ensure the confidentiality, integrity, availability and resilience of the systems and services in connection with the processing in the long term. The state of technology, the implementation costs and the type,



scope, circumstances and purposes of the processing as well as the probability of occurrence and severity of the risk to the rights and freedoms of natural persons, in particular through possible violations of the protection of personal data pursuant to Art. 32, para. 2 GDPR, must be taken into account.

- (3) The technical and organizational measures to be undertaken by the processor are available upon request. Please send an email to privacy@mpi.nl, topic "TOM's IDLAS" to receive the most recent version.
- (4) The processor shall use a procedure pursuant to Art., 32 para. 1 d) GDPR for the regular review of the effectiveness of the technical and organizational measures to ensure the safety of the processing and shall inform the controller of the results of the review.
- (5) During the course of the commission relationship, technical and organizational measures may be modified as part of ongoing technical and organizational changes. Any changes must be set forth in writing.
- (6) Personal or sensitive data may only be transmitted via encrypted connections or in an encrypted form. If the recipient requires a password or other form of key for decryption, this must be transmitted by a different method than the connection to be encrypted. The processor shall bear any costs incurred by them for the encryption.
- (7) Secure communication by e-mail must be carried out using end-to-end encryption. If special categories of personal data within the meaning of Article 9 GDPR and confidential data is communicated by e-mail, S/MIME is to be used. For this purpose, the processor shall ensure during the entire term of the contract that its employees have S/MIME-capable mail clients with signature and encryption function as well as valid certificates which comply with the minimum requirements of the "Notice on electronic signatures according to the Signature Act and the Signature Ordinance (overview of suitable algorithms)" and which are issued by a common certification authority. Outgoing e-mails sent by the processor's employees must be digitally signed as standard. If the recipient requires a password or other form of key for decryption, this must be transmitted by a different method than the connection to be encrypted. The processor shall bear any of its own costs incurred by the encryption.
- (8) Personal or sensitive data must be transmitted via encrypted connections or in encrypted form only. If the recipient requires a password or other form of key material for decryption, this must



be transmitted through a different channel than the connection to be encrypted. The contractor shall bear the costs of encryption incurred on their part.

- (9) Email messages between the parties regarding data protection, secrecy and security will only be accepted if a digital signature is added to the text part of the email. In the case of personal or sensitive content, the message must also be encrypted.
- (10) The processor advises and supports the client in the selection and implementation of the suitable technical and organizational measures in order to effectively implement the data protection principles contained in the GDPR (e.g. data minimization), to incorporate the necessary guarantees into the processing in order to meet the requirements of the regulation and to protect the rights of the persons concerned (privacy by design). The same applies to technical and organizational measures to ensure that only such personal data the processing of which is necessary for the respective specific processing purpose is processed (privacy by default).

Section 7 Controller's Rights and Duties

- (1) The processor shall provide the controller with proof of compliance with the obligations laid down in this contract by suitable means.
- (2) The controller shall have the right to carry out checks in consultation with the processor or to have them carried out by auditors to be named in individual cases, the costs of which shall be borne by the controller. The controller has the right to convince themselves of the processor's compliance with this contract in the processor's business operations by means of spot checks, which as a rule must be notified in good time.
- (3) The processor shall ensure that the controller can satisfy themselves of the processor's compliance with the obligations under Art. 28 GDPR. The processor undertakes to provide the controller with the necessary information on request and in particular to provide evidence of the implementation of the technical and organizational measures.
- (4) The processor shall also be able to provide the following evidence:
 - Certification according to an approved certification procedure in accordance with Art. 42 GDPR;
 - Current certificates, reports or extracts from reports from independent bodies (e.g. auditors, auditors, data protection officer, IT security department, data protection auditors, quality auditors);



- suitable certification through IT security or data protection audits (e.g. according to BSI Grundschutz (German Federal Office for Information Security – Basic Protection);

The controller's rights in accordance with paragraphs 2 and 3 to carry out inspections and to demand information remain unaffected.

- (5) The costs incurred on their side shall be borne by the contracting parties themselves.

Section 8 Subcontracting

(Intentionally left open)

Section 9 Surrender, Deletion, End of Contract

- (1) After complete performance of the contract or upon the controller's request also earlier than that, but at the latest upon termination of the commission, the contractor shall hand over the controller's data as well as the other data stocks in connection with the contractual relationship, all documents obtained by the processor as well as processing and usage results produced shall be handed over to the controller or destroyed in conformity with data protection regulations in accordance with the latter's written instructions. The same applies for test materials and discarded materials. Upon request, the report of deletion or destruction shall be submitted.
- (2) However, backup copies which have been made in order to fulfil liability and warranty claims shall remain unaffected. The processor shall keep these items safe for the controller until they are fully surrendered to the controller and must surrender them to the controller when first asked to do so. The processor shall secure these items against damage or loss in an appropriate manner, in particular by storing and archiving them properly. Backup copies must be locked using suitable methods in order to ensure that the processor cannot use them. The processor shall provide the customer with information on such back-up copies at any time.
- (3) Documents and files which are no longer required must not be destroyed until prior written permission has been given by the controller in accordance with the relevant data protection legislation, providing appropriate evidence.
- (4) Documentation that serves as proof of commissioned processing as well as compliant data processing must be kept by the processor after the end of the contract in accordance with the re-



spective retention periods. At the end of the contract, the controller may hand the documentation over to the controller for their discharge.

Section 10 Liability

- (1) The parties are liable to third parties in accordance with Art. 82 GDPR.
- (2) The internal compensation between controller and processor is based on Art. 82, para. 5 GDPR.
- (3) For fines imposed on one of the parties due to an infringement under Article 83 GDPR, the parties are also liable in internal settlement in accordance with Article 82 para. 5 GDPR.

Section 11 Termination

Each infringement against the provisions of this Agreement shall constitute grounds for extraordinary termination.

Section 12 Final Provisions

- (1) If data of the controller which are held by the processor are jeopardized due to the action of a third party (e.g. execution or seizure), insolvency or settlement proceedings or other events, the processor must immediately notify the controller. The processor must also inform the third party that the data are the controller's data and that the processor processes the data only on behalf of the controller.
- (2) There shall be no retention rights per § 273 Civil Code (*Bürgerliches Gesetzbuch – BGB*) with regard to the processed data or accompanying data media.
- (3) In the event case of contradictions to other contract terms, the provisions of this Agreement have priority.
- (4) Amendments and supplements to this contract and all its components – including any assurances of the contractor – require a written agreement, which can also be made in electronic form, and an express indication that these conditions are to be amended or supplemented. This also applies to the waiver of this form requirement.
- (5) If a provision of this Agreement is invalid, the validity of the other provisions shall remain unaffected. If a provision proves invalid, the parties shall replace it with a new provision which approximates as closely as possible to what the parties intended.
- (6) German law applies.



Date /
Controller signature

Date/
Processor signature

