A New Test of Semantic Association for Use in Awake Craniotomy

A thesis presented in partial fulfilment of the requirements for the degree of

Doctor of Clinical Psychology

at Massey University, Wellington, New Zealand

Jo Chapman

2023

Abstract

This project aimed to develop and validate a new test of non-verbal semantic association to replace the Pyramids and Palm Tree Test (PPTT) in mapping the Inferior Fronto-Occipital Fasciculus (IFOF) during awake craniotomy surgery. Research and clinical experience identified a range of problems with the PPTT, especially in its performance across cultures, and no other existing semantic test met all requirements for effective use in IFOF mapping. A new test was developed based on theoretical and clinical understandings of semantic association, using a novel item format. The final test metrics were an improvement upon the PPTT on all preselected measures of item quality. A total of 707 healthy adult participants were then recruited to complete the new test via an online survey. A final pool of 58 items were selected, all of which performed consistently as intended. Analyses of subgroup performance found there were no clinically significant gender, culture, age, or education effects on scores. These results again suggested the new test improved upon existing tests and achieved its initial design objectives. Validation then proceeded to clinical groups, but recruitment challenges meant too few participants were recruited to permit any analyses. While its validity could not be determined with this study's data, the test and its general population norms will be available for further research and then clinical use if indicated.

Acknowledgements

To all my participants, without whom no research is possible, especially those who generously participated during such a difficult time in their lives.

To my supervisors: Janet Leathem for first believing in me, Stephen Hill for his practical wisdom, and Kwok-Keung Leung for his generosity.

To Mr Kelvin Woon for welcoming me into his operating theatre.

To Morgan Harper for her amazing artwork and her patience with my vague descriptions and requests for changes.

To Shannon Hennig for her advice on supporting people with language impairments.

And to my student colleagues, family, and friends who have so patiently walked alongside me throughout this long journey.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

The idea for this project arose during supervision meetings between Professor Janet Leathem and Dr Kwok-Keung Leung in late 2018. At the time, Dr Leung was working as a Neuropsychologist in the Neurology Department at Wellington Regional Hospital. Part of his role included cognitive assessments of patients undergoing specialist brain tumour surgeries. These operations are performed partly with the patient awake and neuropsychological tests are used to help guide surgeons, showing which parts of the brain can be safely removed and which need to be preserved.

Different tests are needed for each operation depending on the parts of the brain involved. Dr Leung observed that one test was not working well as the test items were out-dated and culturally specific to the United Kingdom. In addition, when he had attempted to use the test in surgery, his patients were consistently unable to perform the task as expected. The reason for this difficulty was unknown. The test is used when tumours are near the Inferior Fronto-Occipital Fasciculus (IFOF), a neural communication pathway which is critical to preserve during surgery. Dr Leung could not find an alternative to the test and was interested in developing a replacement.

Around that time, I was completing an Honours project under Professor Leathem's supervision and had not yet settled on a doctoral thesis topic for the following year. Prior to returning to study, I worked as a software developer with experience in managing complex data, business analysis, and multidisciplinary teams. In terms of my psychology study, I had a strong interest in both neuropsychology and psychometric testing, having completed relevant postgraduate papers. I had also recently enjoyed a guest lecture by Dr Leung, during which he had discussed awake surgery. When Professor Leathem explained the project to me, I felt it was a rare opportunity to be involved in such a specialist application of Clinical Psychology. I also found the concepts intellectually interesting and thought my technology skills could be of value.

Initial meetings with Dr Leung and the Neurosurgeon, Mr Kelvin Woon, confirmed the hospital's need for a new test and the surgical team's willingness to support a doctoral project. The following initial requirements were developed from these discussions:

1. The test should measure semantic association (an aspect of semantic memory).
    a. It should ideally have high specificity to the IFOF.
    b. An 'odd one out' task was proposed, but other types of tasks would be considered if indicated by the literature.
2. The test should be able to be used with a broad range of patients.
3. The test should be able to be used in a surgical context.
    a. It must be quick and easy to administer during surgery.
    b. It should provide immediate feedback to the patient and surgical team.
    c. The test should be able to be delivered on a tablet device.

4. The test could have both verbal and visual forms.
    a. It must be able to be completed by people with a language impairment.
    b. The visual stimuli used need to be free of confounds, such as line density, colour, and familiarity.

The process of validating a new test was discussed, focusing on norming the test in the general population. A key consideration identified was that the validation sample needed to be broad enough to check that the new test worked across cultures. We also discussed comparing scores on the new and old tests.

Following these initial meetings, a three-stage plan was proposed:

1. Develop a new test of semantic association
2. Validate the new test in the general population
3. Trial the new test during surgery

Work began with a review of existing literature to explore relevant theories, test design considerations, and methods. I was also able to observe two awake surgeries to see how current tests are used in context and build an understanding of the surgical process and environment. This thesis will first present the outcomes of that initial work in a literature review, followed by the altered and expanded project plan which was developed from that understanding. Later chapters will detail the methods and results of each stage of the project (the last of which did not proceed as planned). The thesis will conclude with a discussion of the overall outcomes, implications, and future directions.

Chapter 2: Literature Review

This review will present background information about awake craniotomy for glioma, the inferior fronto-occipital fasciculus (IFOF), and semantic memory before discussing existing psychometric tests and their suitability for mapping the IFOF during surgery.

Awake Craniotomy for Glioma

Gliomas are tumours arising from glia, cells within brain tissue that provide a variety of support functions to neurons (Weller et al., 2015). These tumours account for 24% of primary central nervous system cancers and arise most commonly in the frontal and temporal lobes (Davis, 2018).

Glioma are categorised in three ways:

1. by the type of glia involved e.g., astrocytoma (from astrocytes) or oligodendroglioma (from oligodendrocytes)
2. by their origin, either primary (arising in the brain) or metastatic (spread from cancers elsewhere in the body)

3. by their behaviour, using the four tiered tumour grading system developed by the World Health Organisation, ranging from Grade I (benign and slow-growing) to Grade IV (most aggressive) (Louis et al., 2007)

Incidence and mortality vary widely between glioma types, but glioblastoma is the most common and also the most deadly, with a 5-year survival rate of less than 5% (Davis, 2018). In New Zealand, the incidence of high grade (III or IV) gliomas has been estimated at 4.2 per 100,000 people per year for both Māori and non-Māori (Alexander et al., 2010).

The primary treatment for gliomas is generally surgery, with additional radiation and/or chemotherapy for higher grade tumours (Weller et al., 2021). The aim of surgery for gliomas is to remove as much of the tumour as possible (Suarez-Meade et al., 2020). Total tumour removal has been shown to improve survival times over subtotal or partial removal (Brown et al., 2016). However, the extent of tumour removal must be moderated by avoiding permanent severe cognitive deficits due to the loss of functioning brain tissue (Zhang et al., 2020). Tumours of Grade II and above have increasing potential for poorly defined borders and to infiltrate healthy tissue (Louis et al., 2007), making it very difficult for surgeons to determine which tissue needs to, or can, be removed and which does not (Zhang et al., 2020). Scans performed before surgery are not sufficiently precise to define surgical borders and are also progressively less useful during the course of an operation as patients' brains change position in response to the removal of the tumour (Gogos et al., 2020). Further complicating the surgeon's task, the location and size of functional brain regions differ considerably between patients (Vilasboas et al., 2017).

To mitigate these problems, surgeons use awake craniotomy: brain surgery in which the patient is awakened during part of the operation and kept comfortable using local anaesthetic. During an awake craniotomy the patient is asked to perform a task while the surgeon applies direct electrical stimulation to their brain. This allows 'mapping' of brain areas involved in the task based on whether the patient's performance is impaired during stimulation of each area. For example, if the patient is asked to move their arm up and down but is not able to do so when a particular part of their brain is stimulated, then that brain region's function is necessary for arm movement. Tumour removal begins after mapping of the cortical surface is complete and further mapping continues during removal as new areas are exposed (Talacchi, Santini, Casartelli, et al., 2013). Thus, the locations of functional areas can be determined in each individual patient and tumour tissue can be removed precisely up to their borders, with major white matter tracts generally defining the surgical limits (Coello et al., 2013).

The first therapeutic uses of awake craniotomy were in the 1920s for the treatment of epileptic seizures (Penfield & Boldrey, 1937). Then, as now, the goal of surgery was to remove affected areas of the brain while preserving functioning tissue (Penfield & Roberts, 1959) and mapping motor, sensory, and language functions became a well-established technique to support this aim (Whitaker & Ojemann, 1977). Awake surgery with functional mapping was extended to brain tumour treatment in the 1980s (Berger et al., 1989). A meta-analysis in 2012 (De Witt Hamer et al., 2012) found that awake surgery for gliomas produced fewer severe long-term deficits and more complete tumour removal than anaesthetised surgery. Gerritsen et al. (2022) found awake surgery increased both the extent of tumour removal and survival times. Recent meta-analyses of metastatic tumours (Chua et al., 2018) and glioblastoma (Zhang et al.,

2020) found most patients have unchanged or improved neurological functioning after awake surgery and concluded the technique is both safe and effective. The operation is generally well tolerated by patients (Gernsback et al., 2018; Mofatteh et al., 2023) and most potential contraindications (e.g., mental distress, seizures) can now be mitigated (Gogos et al., 2020). It was therefore recommended that awake craniotomy be implemented as the universal standard of care for glioma and it is now considered 'gold standard' (Gogos et al., 2020).

Procedures, equipment, tests, and professionals involved in awake craniotomy vary widely between hospitals and there are no accepted standards (Ruis, 2018; Talacchi, Santini, Casagrande, et al., 2013). A common procedure, and that used by Wellington Regional Hospital, is as follows:

1. Pre-surgical assessment: A neuropsychologist and/or speech language therapist assesses all cognitive domains to establish the patient's current functioning.
2. Surgical planning: Tests/tasks for surgery are selected, considering the brain regions affected by the tumour or the surgeon's planned approach. Test items are filtered to include only those the patient answered correctly in the pre-surgical assessment.
3. Surgery: As each brain region is stimulated, the neuropsychologist or speech language therapist continuously administers the relevant test/task and advises the surgeon of any changes in the patient's performance. They also collaborate with the surgical team to assess the patient's wakefulness and support their comfort.

4. Post-surgical assessment: All cognitive domains are again assessed to determine whether the patient has any new or increased impairment.

Tasks are selected based on the theorised functions of the areas tested and commonly include the patient counting aloud from one to twenty, moving their contralateral arm, and naming objects (Ruis, 2018; Talacchi, Santini, Casartelli, et al., 2013). Intraoperative tasks continue to be researched, developed and updated (Gogos et al., 2020), including experiments with the use of virtual reality and eye-gaze tracking for more complex tasks (Bernard et al., 2023). Some standardised language testing protocols have been proposed (De Witte et al., 2015; Mandonnet et al., 2017) but not widely adopted. The availability and selection of appropriate tasks for each brain region, assessing multiple cognitive domains, is critical to the accuracy of functional mapping (Pallud et al., 2017; Ruis, 2018). Tasks should be simple, short, and applicable across languages, cultures, and literacy levels (Hande et al., 2021).

In summary, the purpose of testing during awake craniotomy is to 'map' brain regions critical to the patient's functioning by repeatedly administering tasks that can normally be completed quickly and easily. If the patient is suddenly unable to respond correctly when a single brain region is deactivated by stimulation, then that region is critical to performing the task and must be preserved. Different tasks are needed for different brain regions, and there are fewer appropriate tasks available for some regions than others. As major white matter tracts often define surgical limits, mapping these is particularly important.

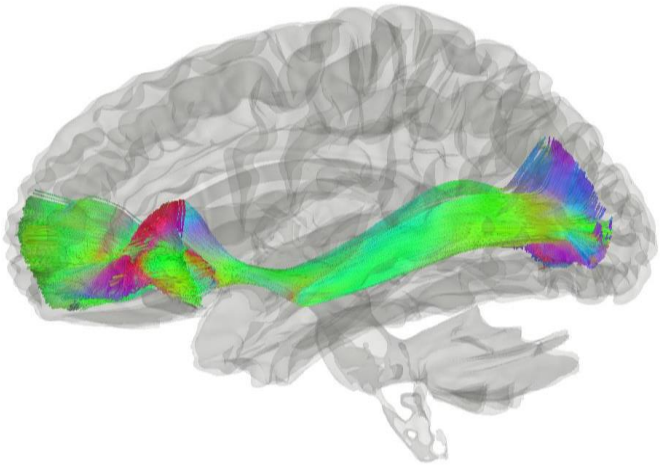
The Inferior Fronto-Occipital Fasciculus

The Inferior Fronto-Occipital Fasciculus (IFOF) is a long white matter tract connecting the frontal lobe of the brain with the occipital and parietal lobes, via the temporal lobe (Bajada et al., 2015). It runs just superior to the Uncinate and Inferior Longitudinal Fasciculi with endpoints in the dorsolateral prefrontal cortex, orbito-frontal area, posterior temporal lobe, dorsal parietal lobe and the occipital cortex (see Figure 1). The IFOF may be a tract unique to human brains as it has not been identified in other primates (Bajada et al., 2015). This suggests it could be involved in distinctly human functions such as language, conceptualisation, and awareness.

Figure 1

The Inferior Fronto-Occipital Fasciculus, based on Diffusion Tensor Imaging scans

From Yeh et al. (2018)



The IFOF is theorised to be a critical part of the ventral pathway for language, known as the 'what' pathway (Duffau et al., 2013). In this theory, language functions are served by both a ventral pathway converting visual information to meaning and a dorsal pathway converting visual information to articulated sounds (Duffau et al., 2014). The IFOF is thought to transfer

auditory and visual information from their relevant cortices directly to the prefrontal control areas to process meaning (Duffau et al., 2013). It has been proposed to consist of a dorsal/superficial layer and a ventral/deep layer, with the dorsal IFOF supporting executive functioning and response selection while the ventral IFOF performs word retrieval, recognition and semantic processing (Rollans & Cummine, 2018).

Given its wide-ranging connections, it is not surprising that the IFOF also appears to be involved in a number of other cognitive processes including goal directed behaviour, emotional functioning, reading, writing, visual perception, and gender identity (Conner et al., 2018; van Heesewijk et al., 2022). Damage to the IFOF is correlated with general cognitive impairments, especially of memory (Chen et al., 2020). Surgical stimulation of the right IFOF has also been found to produce left neglect or rightward deviation in line bisection or target cancellation tasks (Lemée et al., 2018).

Among the IFOF's many functions, the most well researched is its role in semantic memory ('general knowledge' of objects and concepts, discussed in detail in the next section). Measuring the integrity of white matter tracts in people with schizophrenia found that changes to their IFOF were correlated with semantic impairments, while no such correlations were found with other tracts (Surbeck et al., 2020). Han et al. (2013) found that stroke patients with lesions in the left IFOF scored lower on naming and association tasks than patients with lesions elsewhere. They also found that the integrity of the patients' IFOF was correlated with their naming and association scores. Moritz-Gasser et al. (2013) found that stimulating the left IFOF during awake craniotomy produced semantic paraphasias in naming (e.g., saying 'horse' instead of 'zebra') and errors in a non-verbal semantic association task (e.g., choosing a mug to go with

wine instead of a glass). Almairac et al. (2015) measured the infiltration of the left IFOF in patients with low grade glioma and found that increased infiltration was associated with reduced fluency when recalling names of animals but not when recalling words starting with a specified letter (i.e., an impairment specific to semantic memory).

Evidence for the involvement of the right IFOF in semantic memory is less clear. Patients with lesions only in the left temporal lobe perform better on semantic tasks than those with damage on both sides of the brain (Patterson et al., 2015), suggesting the right hemisphere does play a role. In addition, both right and left IFOF integrity are associated with semantic impairments in schizophrenia (Surbeck et al., 2020) and stimulating the right IFOF during surgery can produce non-verbal semantic deficits (Herbet et al., 2017). However, Han et al. (2013) found that right hemisphere lesions were associated with *better* naming performance and suggested this could be because patients with right lesions were unlikely to also have left hemisphere lesions.

As the IFOF is considered 'unresectable' (any surgical damage should be avoided), the tract must be accurately mapped during awake craniotomy (Coello et al., 2013). Mapping is done by repeatedly administering tasks that can normally be completed quickly and easily. If the patient is suddenly unable to respond correctly when a single brain region is deactivated by stimulation, then that region is critical to performing the task and must be preserved. To map the IFOF, surgeons generally rely on semantic memory tasks, such as naming or non-verbal association (Rofes et al., 2015). Semantic memory appears to be the most suitable IFOF function for surgical mapping because it can be quickly tested with few resources, unlike emotional functioning, writing, or complex perception.

Semantic Memory

Human declarative memory is theorised to be divided into episodic and semantic systems (Price et al., 2015). Semantic memory stores 'general knowledge' representations of objects or concepts, including their names, properties, functions, and meaning (Chen et al., 2017). It is required for the use and comprehension of language and is used to recognise objects, infer information about the world, and make predictions. The separation of the semantic memory system from the episodic (memory of events) was first proposed by Tulving (1972). Recent functional imaging research and studies of semantic impairments suggest the two memory systems, rather than being entirely distinct, are based on a common underlying network (Irish & Vatansever, 2020). Brain functions are increasingly being viewed from a 'gradient' perspective, in this case conceptualising memories as existing on a continuum ranging from abstract to concrete time-bound recollections (Irish & Vatansever, 2020). From this perspective, semantic memories are retrieved both as distinct pieces of knowledge and as part of other memories, supporting and expanding a range of cognitive processes.

The characteristics of impairments in semantic memory can provide important insights into its typical functioning. People with such impairments have significant difficulty with language, often asking "What is…?" and using circumlocutory or 'empty' speech such as "that thing over there" (Marshall et al., 2018). They may also find it hard to recognise people or to locate objects, even in their line of sight, and to use objects correctly. While semantic difficulties can have a variety of causes and arise as a symptom of a range of neurological problems, one condition is defined by its selective impairment of semantic memory: the

semantic variant of Primary Progressive Aphasia (svPPA), also called Semantic Dementia (Gorno-Tempini et al., 2011). Individuals with svPPA have atrophy or damage to their anterior temporal lobe which presents as severe difficulties with both naming objects and comprehending single words. They also commonly have difficulty recalling information about objects and reading or writing words with irregular spellings. However, they are able to use correct grammar, articulate speech sounds, and accurately repeat words spoken aloud.

Research into svPPA has identified a number of factors influencing the functioning of semantic memory, including the familiarity of the object and the age at which the name of the object is usually learned (Adlam et al., 2010). Many studies have identified category specific impairments in which individuals appear to lose their knowledge of one category of objects but retain knowledge of another e.g., animals or tools (Chen et al., 2017). The typicality of an object within its category (i.e., sparrows are typical birds while penguins are not) also appears to have an impact. Rogers, Lambon Ralph, Hodges, and Patterson (2004) found patients with svPPA performed as well as healthy participants on tasks where the correct answer was typical of its category but significantly lower than chance when the relationship was reversed. For example, correctly recalling that sparrows can fly but not that penguins can swim. They theorised that knowledge of the general features of categories is preserved longer than knowledge of specific features of individual objects. This theory is further supported by studies of drawings by people with semantic impairments, in which atypical objects are often 'regularised', for example, a camel drawn without its hump or a frog drawn with a tail (Patterson et al., 2007; Rey, 2020).

Theoretical models of semantic memory

To understand the IFOF's potential role in semantic memory, and to inform test design, reviewing theoretical models can be helpful. A number of models have been proposed and include domain-specific, domain-general, and combined views. Domain specific models arose from the findings of impairments in single categories (animals vs tools) or modalities (visual vs sensory) and suggest representations of semantic information are distributed independently throughout the cortex. Most of these models, such as Grounding Representations in Action Perception and Emotion Systems (Martin, 2016), suggest semantic information is stored within or near other relevant cortical structures i.e., visual information in the visual cortex and praxis information in the motor cortex. In contrast, domain general models explain these impairment findings using other factors, such as familiarity or typicality, and propose brain structures dedicated to the storage of semantic information in a general repository (Chen & Rogers, 2014). Many modern models take a combined view, some of which are diagrammed in Figure 2.

One early such view was the Hub and Spoke model (Rogers, Lambon Ralph, Garrard, et al., 2004), in which the Anterior Temporal Lobe (ATL) was proposed to function as a 'semantic hub' connecting a distributed network of concept representations. This 'hub' was argued to support domain-general learning mechanisms which permit abstract concepts and relationships to arise from the collection of domain-specific semantic information. The authors argued the ATL was ideally suited for this purpose due to its wide-ranging connections with many perceptual and motor areas of the brain.

Chen et al. (2017) also argued connections between brain regions, via white matter tracts, were critical to semantic memory. They built a computational model based on scans of

white matter and brain regions activated during semantic tasks, calling it Connectivity Constrained Cognition. This model was able to accurately simulate known semantic disorders and category differences by removing specific connections. Chen et al. argued that semantic functioning should be understood as arising from interactions between a network of interconnected brain regions with domain-general and domain-specific functions.

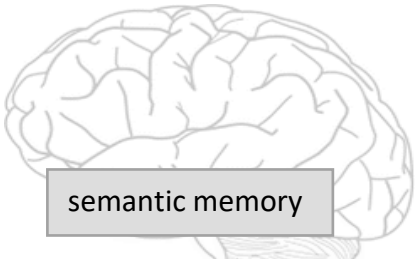
Two other recent models add the idea of a semantic control system serving to coordinate, filter, and focus the retrieval of distributed semantic information. Ralph et al.'s (2017) Controlled Semantic Cognition model extends the hub and spoke model with an integration layer combining the semantic representations from the 'hub' with a representation of the current task. This control network was theorised to interact with, but remain separate from, the semantic representation network. Xu, He, and Bi (2017) propose a Tri-Network Model consisting of experiential, language, and control networks interacting through multiple connector 'hubs'. The experiential network is theorised to support domain-specific information organised by sensory experiences, while the language network supports linguistic and abstract knowledge. In both these models, semantic control networks serve to ensure information retrieved is relevant to the current context or task. Evidence for the existence of a semantic control system comes from patients who make associative errors on semantic tasks, such as saying "milk" in response to a picture of a cow or listing animals as "cat, dog, horse, saddle" (Ralph et al., 2017).

All reviewed models emphasise wide-ranging connectivity between brain regions, in which tracts such as the IFOF may play a part. The Anterior Temporal Lobe is generally considered to play a central role in semantic memory with most models requiring it to be

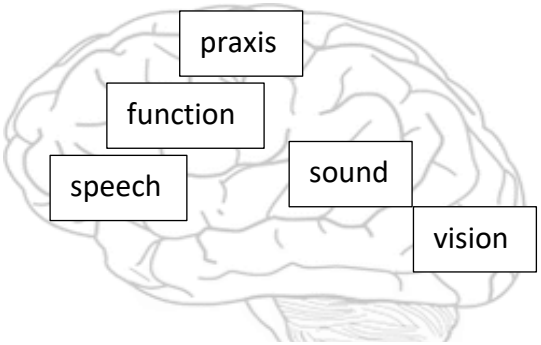
connected to the occipital, parietal, and frontal regions served by the IFOF. The tract may support the integration of information from different regions or may form a part of a semantic control system. In either case, the IFOF can be theorised as necessary for the performance of tasks requiring complex integration of semantic knowledge.

Figure 2

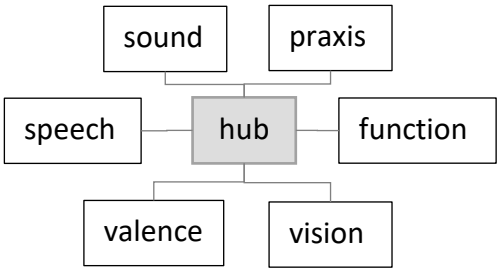
Models of semantic memory



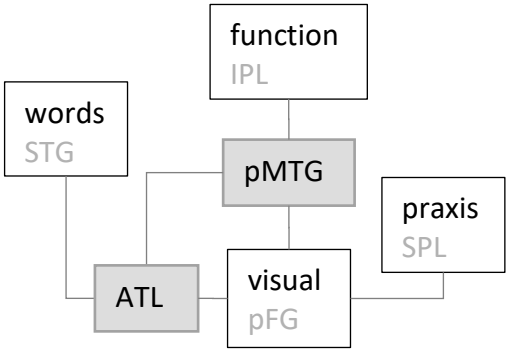
Domain-general model



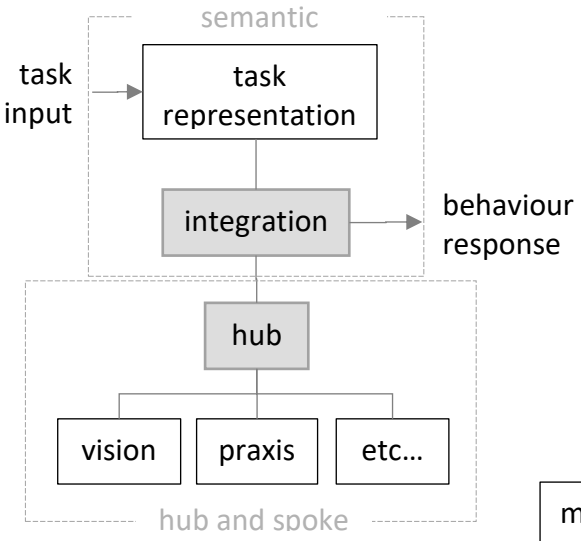
Domain-specific model



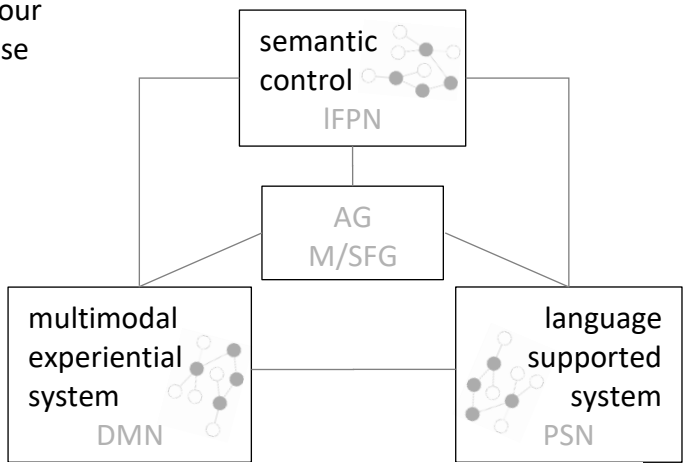
Hub and Spoke model
Rogers, Lambon Ralph, Garrard, et al., (2004)



Connectivity Constrained Cognition
Chen et al. (2017)
ATL: Anterior Temporal Lobe, IPL: Inferior Parietal Lobule, pFG: Posterior Fusiform Gyrus, pMTG: Posterior Middle Temporal Gyrus, SPL: Superior Parietal Lobe, STG: Superior Temporal Gyrus



Controlled Semantic Cognition
Ralph et al. (2017)



Tri-Network Model
Xu et al. (2017)
AG: Angular Gyrus, DMN: Default Mode Network, lFPN: left Fronto-Parietal Network, M/SFG: Middle and Superior Frontal Gyri, PSN: PeriSylvian Network

Testing Semantic Memory

A wide variety of semantic memory tests are used for both research and clinical purposes such as diagnosis of semantic impairment and monitoring the progress of disease as well as mapping in awake craniotomy. Several currently available tests will be described in turn, before discussing the specific test features relevant for the purpose of mapping the IFOF.

Naming and vocabulary tests

The simplest and most common tests of semantic memory are naming tasks, such as the Boston Naming Test (Kaplan et al., 1983), in which the examinee is presented with an image and asked to name the object depicted. Naming has been described as the strongest basis for distinguishing the semantic variant from other forms of PPA (Stockbridge et al., 2021). Various naming tasks are widely used in awake craniotomy to map language areas (Talacchi, Santini, Casartelli, et al., 2013), including some designed specifically for the purpose (Ohlerth et al., 2019). A carrier phrase (e.g., “This is a”) is usually added to differentiate errors arising from various language-related brain functions (see Table 1). Vocabulary and word-to-picture matching tests also measure semantic memory in a similar way. These tests are all vulnerable to language impairment as they generally require spoken responses and/or written stimuli. Further, they ask the examinee to access only one piece of semantic information and thus may not require the activation of semantic control and integration systems using the IFOF.

Table 1

Example of the identification of brain functions using a naming task

Test item	Patient's response during stimulation	Function of stimulated brain region
	Correct: "This is a dog"	None relevant
	Anomia: "This is a… uh…"	Semantic memory or word retrieval
	Semantic paraphasia: "This is a cat"	Semantic memory
	Articulation difficulty: "Diff if a dock"	Oral movements
	No response	Speech production

One variant on naming is the picture word interference task (PWI), in which the image of the object to be named is presented along with a 'distractor' word (e.g., a cat with the word 'horse' superimposed). In research, semantically-related distractor words increase the time taken to name target objects significantly more than unrelated distractor words (Ries et al., 2019). This effect is theorised to be due to the over-activation of similar alternative concepts. PWI tasks have been used to map the IFOF in awake craniotomy (Gogos et al., 2020), theorising that they require the activation of semantic control systems to manage the interference (Ries et al., 2019).

Category fluency

Verbal fluency tasks involve asking examinees to say as many words as they can in one minute, following some rules, and are generally used to assess executive functioning (Strauss et al., 2006). However, category fluency, in which words must be from a given category (e.g.,

animals), is also regarded as a valid and sensitive test of semantic memory across a range of populations (Adlam et al., 2010; Neill et al., 2014). Factor analyses by Schmidt et al. (2017) suggest category fluency assesses cognitive functions which are distinct from those assessed by letter fluency (words starting with a given letter). These two fluency scores can therefore be compared to determine whether an examinee's impairment is specific to semantic memory or related to broader language and/or executive functioning (Henry & Crawford, 2004). Category fluency has been described as one of the most well-researched, validated, and understood neuropsychological tests (Ardila et al., 2006) and norms for both letter and category fluency are available for a wide range of populations (Strauss et al., 2006). However, due to their length, fluency tests are impractical for use during awake craniotomy. Significant language or executive function impairments can also render the task so difficult for examinees that score comparisons are not meaningful.

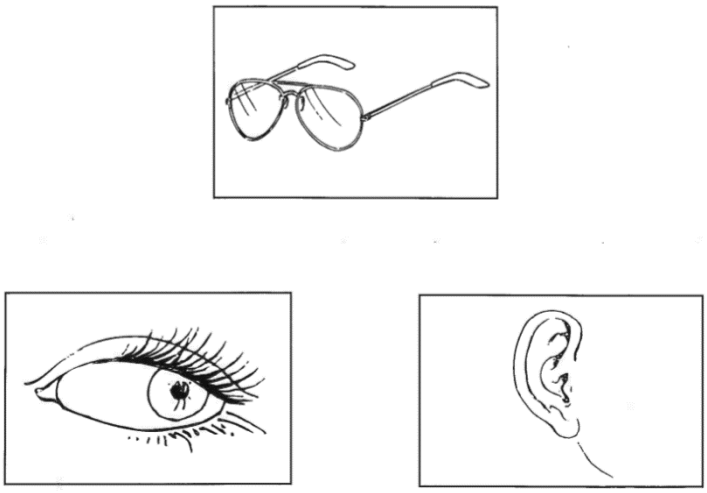
The Pyramids and Palm Trees Test

The Pyramids and Palm Trees Test (PPTT) is a test of semantic association widely used in research (Olivarez & Boroda, 2007) and to map the IFOF during awake craniotomy, both in New Zealand and elsewhere (Chang et al., 2018). The PPTT was developed by Howard and Patterson (1992) with the aim of assessing an individual's ability to access detailed semantic information but not specifically for use in awake craniotomy. Each item consists of three stimuli (images or words) and the examinee is asked which of the bottom two objects "goes with" the object presented at the top. An example item is shown in Figure 3. The PPTT was initially trialled in the United Kingdom with healthy adults and a small group of young adults in rehabilitation for non-head-related traumatic injury. All of their participants scored in the range of 49-52 correct

out of 52 test items, showing a clear ceiling effect. The authors state a score of 35 is better than chance at $p<.01$ but also say score interpretation should be based on the 'overall pattern' of performance between picture, word, and oral versions of the test (i.e., whether some scores were better than others).

Figure 3

An example item from the PPTT, to which the correct answer is the eye



A Mental Measurements Yearbook review of the PPTT found that, despite being widely used, the test had no clear theoretical background or definition of the construct being tested and that the images used were of variable quality and depicted out of date items (Olivarez & Boroda, 2007). Fourteen years later, these items are even further out of date. International studies consistently find a correlation between PPTT scores and years of formal education (Callahan et al., 2010; Gamboz et al., 2009; Gudayol-Ferré et al., 2008), in one case explaining 20% of the variance in scores (Bozdemir & Gurvit, 2021). This raises concerns about its validity

for measuring impairment in some populations. Some studies also found a correlation with age, although this is difficult to separate from the education effect as older participants were generally also less educated (Callahan et al., 2010; Gamboz et al., 2009). An item response theory analysis of the PPTT concluded it is only suitable for measuring semantic ability at extreme low functioning (Fergadiotis et al., 2010).

The PPTT is also less valid with individuals from cultural backgrounds outside of the United Kingdom. Mean scores of healthy adults from some countries fall outside the published normal range of 49-52 (see Table 2). Considerable differences in cultural relevance between items on the PPTT can be inferred both from item analyses in previous studies and from the PPTT adapted versions made in China (Guo et al., 2014), Spain (Martínez-Cuitiño & Barreyro, 2010), Quebec (Callahan et al., 2010), and the United States (Klein & Buchanan, 2009). Notably, two of the 52 items relate to British-style circuses, three to Christianity, and another two to stereotypes about indigenous Arctic peoples (using the pejorative term 'eskimo' in the word version). Two PPTT items (windmill goes with tulip and acorn goes with pig) are removed from all adapted versions, suggesting they do not function at all across cultures, while a further eight items are kept only in studies which made minimal alterations to the original test. In contrast, Breining et al. (2015) selected 14 items from the PPTT for a brief version of the test and these 14 items are also kept in every cultural adaptation reviewed for this study, suggesting they perform consistently in a variety of populations. The selection criteria for these 14 items was that the associations were based on the defining features of the objects, rather than cultural knowledge, and the pictures were clear and easy to interpret.

Table 2

Mean Pyramids and Palm Trees Test Scores in International Studies

Study	Sample	Mean score (SD)*
Callahan et al. (2010)	214 Quebec-French adults	49.3 (1.98)
Gamboz et al. (2009)	464 Italian adults aged 49+	47.52 (5.24)
Klein and Buchanan (2009)	90 US college students	47.9 (1.97)
Mehri et al. (2018)	270 Iranian adults	49.26 (2.67)
Gudayol-Ferré et al. (2008)	234 Spanish adults	51.1 (1.3)
Rami et al. (2008)	121 Spanish adults aged 60+	49.6 (2.2)
Breining et al. (2015)	20 Argentinian adults	49.8 (2.63)
	12 Greek adults	47.6 (2.07)
	18 US adults	49.4 (1.46)
Bozdemir and Gurvit (2021)	194 Turkish adults	48.88 (5.72)

*The published normal range is 49-52

Semantic test batteries

Batteries of semantic tests have also been developed, aiming to assess multiple aspects of semantic memory through a variety of tasks. The Cambridge Semantic Memory Test Battery consists of naming, sorting, word-picture matching, and association tasks, all using the same set of category-balanced items (Adlam et al., 2010). Subsets of items matched on their familiarity or age of acquisition are also provided. The association task within the Cambridge battery is known as the Camel and Cactus Test (CCT). It uses the same format as the PPTT but has four response options for each item, intended to render the test more sensitive to subtle semantic impairments (see Figure 4 for an example). Adlam et al. (2010) report participants with svPPA scored lower than control participants on all battery tasks in the validation study, although the

battery did not clearly distinguish participants with svPPA from those with more general cognitive impairment. While the balancing of items in this battery is a useful advance on other tests, card sorting tasks are not practical for use in surgery (requiring significant space, physical manipulation, and time) and the CCT contains many culturally specific items (e.g., peacock goes with hat and strawberry goes with tennis). Despite this, it has been successfully adapted and validated for use in India (Paplikar et al., 2022).

Figure 4

Example item from the Camel and Cactus Test, to which the correct answer is the hammer

Another battery was recently developed by Ohman et al. (2020) and is intended as a screening tool for semantic impairment. Tasks in the battery involve naming, matching, answering factual questions, and identifying common features of objects. Ohman et al. (2020) report participants with cognitive impairments received lower scores than healthy controls,

scores on the battery were positively correlated with scores on existing semantic tests, and the reliability of the battery was high on all measures. However, their interviews with clinicians identified potential educational and cultural bias in the test items, for example “Does a piano have strings?”.

Other tests

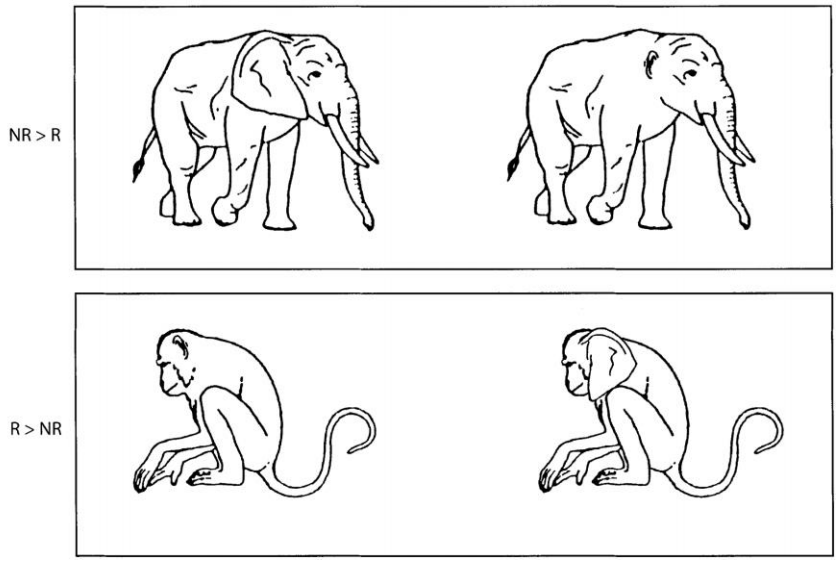
Research into svPPA has used all of the above tests, as well as word-nonword decision tests in which the examinee is asked to distinguish real words from pseudo-word distractors (Rogers, Lambon Ralph, Hodges, et al., 2004). Other researchers have used a similar format to the PPTT, but with small alterations. The Kissing and Dancing Test is one such variant, using verbs instead of nouns (Baradel et al., 2014).

Two visual tests developed specifically to study svPPA were reviewed, both of which reported their design process in detail. The Over-regular Object Test (OOT) presents examinees with two images: an object and a pseudo object with a single feature altered e.g., an elephant with small ears (Rogers, Lambon Ralph, Hodges, et al., 2004). The examinee is then asked to “point to the real thing”. The features altered are balanced between those which are typical of the object’s category and those which are not (see Figure 5 for examples). The Feature Reality Test (FRT) similarly asks examinees to point to the “more real looking” of two images (Garrard & Carroll, 2006), with the objects being altered by colour (an orange banana), context (a horse in a shopping mall), or motion (a rearing cow). Both plausible and implausible distractors are used e.g., an orange banana and a pink banana. These two tests were able to investigate specific features of svPPA, including the severity and progression of impairments over time, but are too narrowly focused for use in mapping a complex associative tract such as the IFOF.

Figure 5

Example items from the Feature Reality Test, first where the correct answer is atypical of the object's category and second where the relationship is reversed

From Rogers, Lambon Ralph, Hodges et al. (2004)



Further tests of semantic memory were developed during the period of this study and are briefly summarised here:

- Le Test de Dessin: measures errors in participants' drawings of common objects (Rey, 2020)
- The Brief drawing task: measures errors in drawings of familiar and less familiar objects (Pozueta et al., 2019)
- The 42-Item Semantic Memory Test (SMT-42): verbal questions about semantic features of objects e.g., does a rabbit have feathers or fur? (Mazoué et al., 2022)

- A task using the same format as the PPTT but able to discriminate different types of semantic associations (Luzzatti et al., 2020)
- The Verst-Maldaun Language Assessment (VMLA): naming and semantic tasks using the same format as the PPTT and designed for use during awake craniotomy in Brazil (Verst et al., 2021)
- The Ikos Test: uses the same format as the PPTT with an additional distractor image (choice of 3) and stimuli that are everyday objects associated by their function, designed for the diagnosis of Alzheimer's Disease (Cejudo et al., 2022)
- The Free Association Task (F-assoc): compares examinees' responses to cue words with those given by healthy adults, designed for use in conjunction with verbal fluency to differentiate semantic impairment from executive function deficits (Zannino et al., 2021)

Using semantic tests for IFOF mapping

When selecting semantic tests for the purpose of mapping the IFOF during awake craniotomy, some specific test characteristics should be considered. These characteristics were identified through discussions with the surgical and research team, described in the Introduction, and were also apparent as themes throughout the literature reviewed.

1. *Evidence for validity*: as in any clinical context, there should be evidence that a test is a valid measure of semantic memory.
2. *Language independence:* the test should ideally be independent of language to allow IFOF mapping in people with language impairments (either due to their

tumour or to the surgery itself). Language independent tests are also more likely to be able to map the IFOF in the right hemisphere. While all tests are likely to depend on language to communicate their initial instructions, responding correctly to an individual test item should not require an examinee to understand or produce words in either written or oral form.

3. *Association:* a test for IFOF mapping should be 'associative', in this context meaning that it requires the integration of multiple pieces of semantic information. This consideration was informed by the semantic memory models reviewed, which suggest tasks involving integration and control of semantic information retrieval are more likely to activate the IFOF.
4. *Cultural fairness:* a test for use in surgery should be able to be validly used with people from a range of cultural and educational backgrounds. Tumours do not discriminate, but their treatment may if appropriate tests are not available for some patients.
5. *Speed to administer:* a test used for mapping should be quick to administer because electrical stimulation can only be applied to the brain for a few seconds at a time.

Table 3 summarises these characteristics for each semantic test discussed previously (batteries of tests are not included as they are evidently impractical in a surgical context). Currently, naming, PWI, and the PPTT are commonly used for IFOF mapping but, as shown in the table, neither these nor any other existing tests are entirely suitable for this purpose.

Table 3

Summary of features of existing semantic memory tests relevant to IFOF mapping

	Evidence for validity	Language independent	Associative	Culturally fair	Quick to administer
Naming	✓	✗	✗	?	✓
PWI	✓	✗	?	?	✓
Category fluency	✓	✗	✗	✓	✗
PPTT / CCT / VMLA	?	✓	✓	✗	✓
FRT / OOT	✓	✓	✗	?	✓
SMT-42	✓	✗	✗	?	✓
Drawing	✓	✗	✗	?	✗

PWI: Picture Word Interference
PPTT: Pyramids and Palm Trees Test
CCT: Camel and Cactus Test
VMLA: Verst-Maldaun Language Assessment
FRT: Feature Reality Test
OOT: Over-regular Object Test
SMT-42: 42-Item Semantic Memory Test

Conclusion

Awake craniotomy allows for safe and effective glioma removal, but it relies on the accuracy of the intraoperative tests used for brain mapping. Brain regions critical to the patient's functioning are identified by repeatedly administering tasks that can normally be completed quickly and easily. If the patient is suddenly unable to respond correctly when a single brain region is deactivated by stimulation, then that region is critical to performing the task. White matter tracts perform important functions and must be preserved; thus, they often form the limits of surgery for glioma removal. The IFOF is one such tract involved in a wide variety of functions, especially semantic memory. Impairment in semantic memory can be disabling as it is required to use and understand both language and objects in the environment.

Mapping the IFOF during awake craniotomy is therefore critical to preserving this ability after surgery. The tests most commonly used to map the IFOF are naming and the Pyramids and Palm Trees Test (PPTT). Both approaches have problems, raising concerns about the accuracy of mapping currently possible and therefore the risk of IFOF damage during surgery. Reviews of existing semantic tests did not identify any better suited for this purpose, although they have contributed to our understanding of specific patterns in semantic impairment.

This review confirmed the need for a new test designed specifically for use in awake craniotomy, able to be used with patients from a wide variety of backgrounds, and both specific and sensitive to IFOF impairment. Such a test would improve the guidance available to neurosurgeons and therefore outcomes for glioma patients. Non-verbal semantic association is the most viable IFOF function to use for mapping as it is well-researched and possible to quickly test in a surgical context. Commonalities between current models of semantic memory support the IFOF's likely role in semantic control and integration of information. The sensitivity of the Pyramids and Palm Trees Test to IFOF stimulation in some populations also indicates non-verbal semantic association can be a valid mapping tool.

Chapter 3: Project Outline

Based on the literature review, surgical observations, and discussions, a more refined project plan was developed. The plan included four phases which are summarised below. Detailed methods and results are reported for each completed phase in following chapters.

Phase 1 – Test Development

Objective: Develop a new test of non-verbal semantic association

Key Questions: Are new test items of higher quality than PPTT items?

- Intended outcome: yes

Tasks: Determine test design principles and priorities

Select item quality measures

Select an item format

Generate test items

Produce image stimuli

Calculate item quality measures and compare with the PPTT

Phase 2 – General Population Validation

Objective: Determine whether the test can be completed by healthy adults from a variety of backgrounds

Key Questions: Do all test items perform similarly?

- Intended outcome: yes

Can most people score near ceiling on the new test?

- Intended outcome: yes

Do age, education, or culture influence scores on the new test?

- Intended outcome: no

Tasks: Design an online survey to administer the new test

Recruit a large and varied sample of healthy adults to complete the survey

Analyse the relative performance of test items

Refine the item set

Compare test performance between groups

Phase 3 – Clinical Group Validation

Objective: Determine whether the test is a valid measure of semantic and/or IFOF impairment

Key Questions: Are scores on the new test related to scores on existing semantic tests?

- Intended outcome: yes, similar or positively correlated scores

Do people with semantic impairment score lower than the general population?

- Intended outcome: yes

Do people with glioma score lower than the general population?

- Either outcome possible

Does IFOF impairment influence scores on the new test?

- Intended outcome: IFOF impairment related to lower test scores

Tasks: Select comparison tests
Recruit people with glioma to complete the tests
Recruit people with semantic impairments to complete the tests
Compare new and existing test scores
Compare test performance between groups

Note: While this phase was conducted as planned, the obtained sample size was much smaller than anticipated. As a result, the main objective of Phase 3 was not achieved. This will be discussed further in the relevant chapter.

Phase 4 – Trial in Awake Craniotomy

Objective: Determine whether the test is practical for IFOF mapping in awake craniotomy

Key Questions: Can patients complete test items during surgery?
- Intended outcome: yes

Does stimulation of the IFOF impair test item performance?
- Intended outcome: yes

Tasks: Plan trials with the surgical team
Conduct trials

Note: Due to an increase in the size and complexity of earlier phases, the time constraints of a DClinPsych project, and a reduction in the frequency of awake surgeries performed at Wellington Regional Hospital, it was later decided to remove this fourth phase from the project scope. It may be conducted in future by another student or by hospital staff.

Chapter 4: Test Development

This chapter discusses the methods used and outcomes of the initial development of a new semantic association test. General design principles will be explained, followed by the selection of test format, quality measures, items, and images. Finally, the characteristics of the resulting new test will be discussed in comparison to the PPTT.

Design Principles

The Standards for Educational and Psychological Testing (American Educational Research Association et al., 2014) define three foundational principles for the design of any test:

1. ***Validity*** includes both evidence of accuracy and the specification of the test's intended use, interpretation, and theoretical and/or empirical basis. A valid test has a clear purpose and has been shown to fulfil that purpose. During the design of this test, key validity considerations were the test's appropriateness for use during surgery and its sensitivity and specificity to IFOF or semantic impairment.
2. ***Reliability***, or precision, relates to the testing procedure, scoring, and errors of measurement. A reliable test represents the measured construct without being

unduly influenced by other factors. The simplicity and clarity of test instructions and stimuli was a priority for this test, as well as the avoidance of unnecessary demands on an examinee during surgery.

3. ***Fairness*** includes the absence of bias as well as a test's accessibility. A fair test treats all test takers equitably and is valid for all members of the intended population. As discussed in the Literature Review, fairness was a key principle informing the need for a new test. The intended population for this test was all people with gliomas, regardless of their age, ethnicity, country of residence, or educational background.

This test was developed specifically to detect impairment during surgery and not to measure individual differences in semantic ability. The purpose of testing in awake surgery is to 'map' brain regions critical to the patient's functioning by repeatedly administering tasks that can normally be completed quickly and easily. If the patient is suddenly unable to respond correctly when a single brain region is deactivated by stimulation, then that region is critical to performing the task and must be preserved. This means that some aspects of traditional test design were considered differently. Notably, all items in the test were intended to be of similar difficulty and measure the same functions. This allows performances on different items during stimulation of different brain regions to be directly compared with each other but is contrary to the usual aim of avoiding overrepresentation of concepts (Haynes et al., 1995). Test design also usually aims to avoid items answered correctly by most examinees (Haladyna, 2015) but this test was intended to be normed at ceiling so that any failed items are indications of impairment. This also allows all items to be used during surgery without restricting the item

pool by needing to exclude items patients were not able to answer correctly in a presurgical assessment.

Further specific priorities for the test design arose from the characteristics used in the Literature Review to evaluate existing semantic tests: language independence, association, and speed to administer. The ease of test administration during surgery, for both staff and patients, was also considered throughout.

Theoretical basis

No single semantic model was selected as a theoretical basis for this project. Rather, the common elements across models were considered to provide sufficient grounds for test development. All reviewed models (Chen et al., 2017; Ralph et al., 2017; Rogers, Lambon Ralph, Garrard, et al., 2004; Xu et al., 2017) suggest connectivity between brain regions is critical to semantic functioning, and a likely role of the IFOF in integrating and/or controlling the retrieval of semantic information. Tasks requiring complex integration of semantic knowledge can thus be theorised to be sensitive to IFOF impairment. Studies also suggest non-verbal semantic tasks involve both sides of the brain (Herbet et al., 2017; Patterson et al., 2007; Surbeck et al., 2020), while more language-based tasks are largely performed by the left or dominant hemisphere (Almairac et al., 2015; Han et al., 2013; Moritz-Gasser et al., 2013). Based on these findings, non-verbal tasks were deemed more suitable to allow IFOF mapping on either side of the brain.

Test Format

A key consideration in the choice of format was the resulting simplicity of the task in terms of its instructions, the process of determining an answer, and the physical response required. The effects of the question format on the likely sensitivity and specificity of the test were also considered, as well as the ease of adding response options to reduce chance effects. As category knowledge appears to be preserved longer than specific knowledge (Patterson et al., 2007), tests requiring specific knowledge were considered more likely to be sensitive to mild semantic impairment. Specificity considerations were less clear as the IFOF is a complex tract performing many functions. Thus, tests additionally requiring some executive functions are less specific to semantic impairment but could be more sensitive to IFOF impairment.

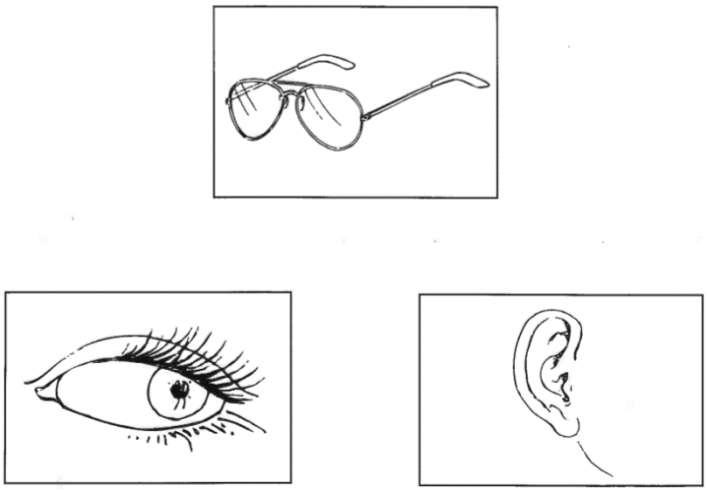
Three potential question formats for the new test were reviewed: single probe and targets, odd one out, and composite images.

Single probe and targets

This is the format used by the Pyramids and Palm Trees Test (PPTT), the Camel and Cactus test from the Cambridge battery, and others. It consists of a single image at the top of the page, with other images underneath. The examinee is asked which of the bottom images "goes with" the one at the top. An item from the PPTT is shown in Figure 6.

Figure 6

An example item from the PPTT, to which the correct answer is the eye



This format allows all target items to be from the same category (i.e., the eye and ear are both body parts) thus requiring the examinee to differentiate their specific features in order to answer correctly. The examinee knows that the single image at the top must be part of the target set, so planning requirements are minimised as each image beneath can simply be examined in turn. To reduce the effects of chance, more potential targets can be added without significantly increasing the complexity of the task, as in the Camel and Cactus Test, which uses four (Adlam et al., 2010).

To check that the examinee understands the task, some psychologists ask them to point to both the probe and the selected target, thus demonstrating that "these two go together" (K.-K. Leung, personal communication, December 17, 2018). This double-point response increases the demands on the patient in a surgical context. Further, the reports of patients unable to complete the PPTT during awake craniotomy at Wellington Regional Hospital suggest there are

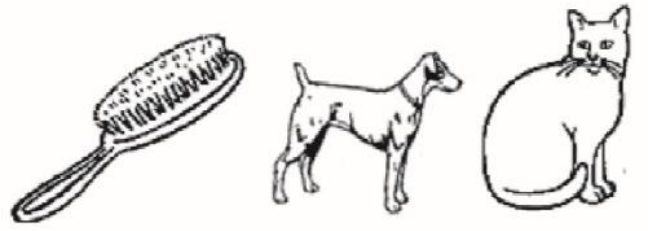
further unidentified demands of this task which exceed the capacity of many patients during surgery.

Odd one out

This is the format originally proposed for the new test by Wellington Regional Hospital staff. It involves three images, of which the examinee is asked to identify the "odd one out" by finding an association between the other two. Generally, the "odd one out" is from a different category than the other objects. Figure 7 shows an example, in which the examinee would be expected to point to the hairbrush as it is not an animal.

Figure 7

Example item from the Dutch Linguistic Intraoperative Protocol (De Witte et al., 2015)



The examinee is only required to point to a single target and the task is widely used in children's activities, meaning it may be familiar to examinees and thus easy to understand. Executive functioning is required to identify the target set as examinees generate potential associations, test theories, discard those that do not produce an answer, and generate new options. While less specific to semantic impairment, these demands are likely to involve the frontal brain areas served by the IFOF.

However, this format best lends itself to associations based on category membership, rather than usage, location, or other associations. Targets and distractors cannot be from the

same category, otherwise they would create a potential set that is not the one intended. This means examinees do not need to differentiate specific features in order to answer correctly. As discussed earlier, using only category knowledge is likely to render this task much less sensitive to semantic impairment.

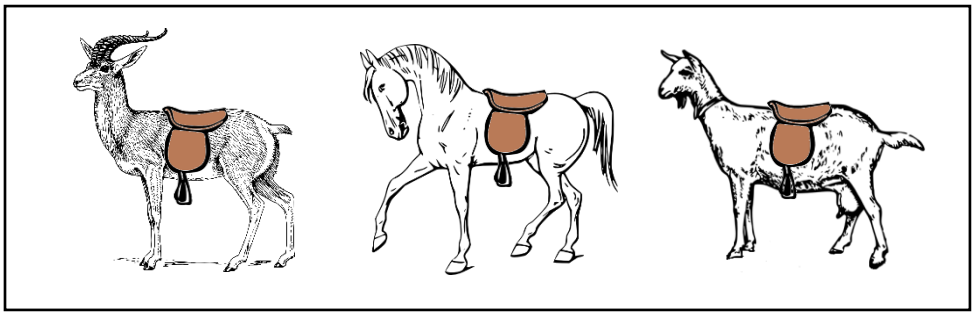
Testing each potential combination of items to find the set involves significant planning (i.e., check if items 1 & 2 go together, then items 2 & 3, then items 1 & 3). If more items are added to reduce chance effects, the complexity of planning required increases exponentially. Further, responding correctly requires the examinee first to identify the correct set and then point to the *other* item instead. These additional demands could be confounding factors if a patient's executive functioning, working memory, or response inhibition is affected by their condition or the surgical context.

Composite images

This is a novel test format conceived by the author specifically for this project based on the designs of the Feature Reality Test (Garrard & Carroll, 2006) and the Over-regular Object Test (Rogers, Lambon Ralph, Hodges, et al., 2004). Associations are depicted by merging the objects into one composite image for each potential set and the examinee is asked to identify "which picture is correct". Figure 8 shows an early demonstration of the format, based on item 5 from the PPTT, in which the examinee would be expected to point to the saddled horse.

Figure 8

Early demonstration of the composite images test item format



The composite images format was intended to reduce the demands on a patient in the surgical context by simplifying the task's instructions, concept, and response. It allows distractor objects to be from the same category as targets and thus requires examinees to differentiate specific features. This format also permits adding more distractors to reduce chance effects without significantly increasing the complexity of the task. It places only minimal demands on an examinee's planning and executive function abilities.

However, correct performance of this novel task could depend more on identifying absurdity than on making semantic associations. The reduced executive functioning requirements could also reduce the test's ability to detect IFOF impairment. Further, this format restricts the potential item set, as some associations cannot be easily depicted in a composite image, and it is not possible to produce a parallel word version of the test.

Format selection

The potential formats reviewed are summarised in Table 4. All formats required only pointing responses but differed in their demands on the examinee's planning, working memory, and other executive functions. In terms of sensitivity to semantic impairment, test formats

differed mainly in whether examinees were required to differentiate specific features of objects. Despite some potential disadvantages, the composite images format was deemed more likely fit for purpose than either of the other formats considered. Thus, test development proceeded using this format.

Table 4

Summary of test format considerations

	Probe and targets	Odd one out	Composite images
Simplicity of instructions	Moderate	Moderate-High	High
Simplicity of task	Low	Low	High
Likely sensitivity	High	Low	Moderate-High
Likely specificity	High	Moderate	Moderate-High
Ability to add response options	High	Low	High

Quality Measures

Semantic test quality measures were selected before test development began, based on the literature and design priorities discussed. They were then used to guide test development and to compare the new test with the PPTT once it was complete. The new test aimed to show an improvement over the PPTT in all areas of quality selected for measurement. These included semantic category representation, age of word acquisition, word familiarity, and strength of association.

Semantic categories

Given the well-documented occurrence of category-specific semantic impairments, test items should represent objects from a range of semantic categories in order to detect impairment in any area. Ideally, all categories should be evenly represented to avoid biasing

the test towards particular types of impairment. No standardised semantic category lists were identified, so this study began with those used in the Cambridge Semantic Battery: animals, birds, fruit, household items, tools, and vehicles (Adlam et al., 2010). These categories were used to begin classifying the objects in PPTT test items and were then gradually added to and combined until all PPTT items were able to be categorised and categories felt distinct. The final categories used for this study are listed in Table 5, with examples taken from PPTT items.

Table 5

Semantic categories used during test development

Category	Examples
Household item	Pencil, Lamp, Ticket
Food	Apple, Milk, Egg
Animal	Dog, Butterfly, Owl
Nature	Tulip, Moon, Rain
Building	Window, Church, Kennel
Furniture	Desk, Bed, Curtain
Clothing	Boots, Ring, Waistcoat
Body part	Ear, Hands, Heart
Vehicle	Bicycle, Bus, Rowboat
Person	Baby, Soldier, Girl
Tool	Hammer, Screw, Saw

Age of acquisition and familiarity

Age of word acquisition and word familiarity are psycholinguistic measures which have both been shown to be related to performance on semantic tasks (Adlam et al., 2010). In order for test items to be of comparable difficulty, the age of acquisition and familiarity of the names of objects used should be similar. Further, knowledge of objects with a low age of acquisition and high familiarity seems less likely to be affected by education. Test items using such objects may therefore be fairer for examinees across a range of levels of education. Many tables of

psycholinguistic normative data have been published and the Glasgow norms (Scott et al., 2018) were chosen for this study. They were the most recent data set reviewed and contain both age of acquisition and familiarity ratings reported on similar scales (shown in Table 6).

Table 6

Selected rating scales used in the Glasgow norms (Scott et al., 2018)

Rating	Familiarity *How often seen/heard, how recognisable*	Age of acquisition *Age when first learned*
1	Very unfamiliar	0-2 years
2		3-4 years
3		5-6 years
4	Moderately familiar	7-8 years
5		9-10 years
6		11-12 years
7	Very familiar	13+ years

Strength of association

The strength of association between pairs of words can be measured using Latent Semantic Analysis (LSA), a theory and mathematical tool for modelling semantic relationships and language learning (Landauer & Dutnais, 1997). LSA-based mathematical models are built from a corpus, or selection of texts. Once a model has been built, numeric ratings (from 0-1) of the strength of association between pairs of words can be obtained e.g., “dog” and “puppy” have a score of 0.76, while “dog” and “wool” have a score of 0.22. In a semantic test, high-strength associations between objects in target sets suggest that the test developers’ ideas of sets align with those of other people, while low-strength associations between targets and distractors suggests a reasonable choice of distractor items without creating unintended alternative sets. For every item, the associative strength of the target set should be higher than

any associations with distractors. Association strengths should also be similar across items if they are of comparable difficulty. For this study, a freely available model built from all Wikipedia articles using the 'GloVe' method was chosen because it had the best performance on the Google Analogies semantic association benchmark test (Fares et al., 2017). The Python package 'Gensim' (Řehůřek & Sojka, 2010) was used to interact with the model, using scripts written and run in Microsoft Visual Studio Code.

Item Selection

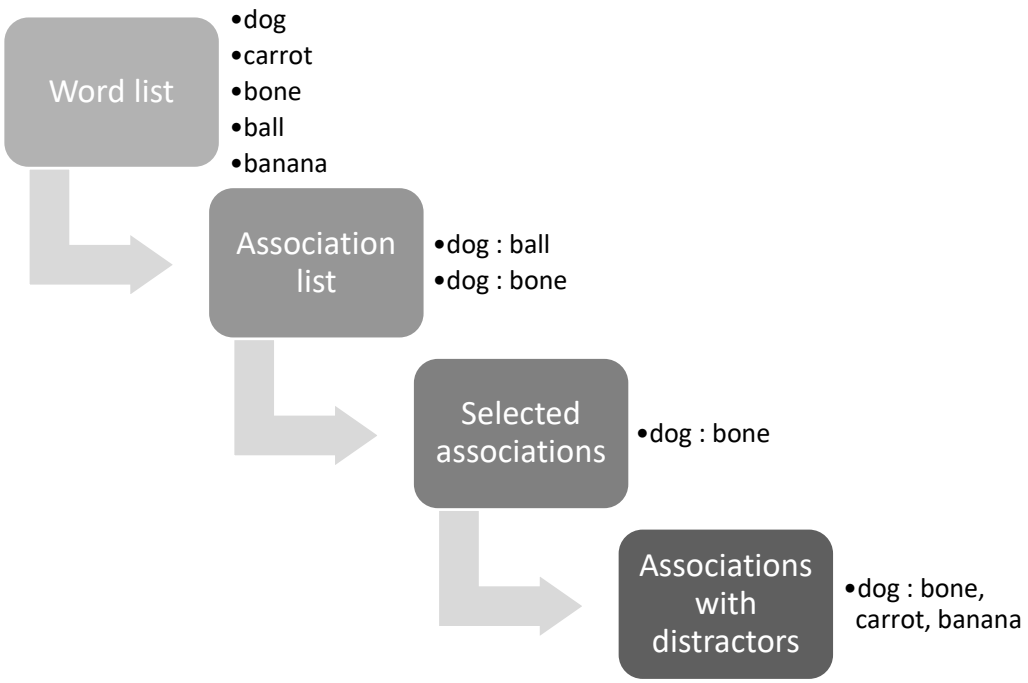
Producing the item pool for the new test proceeded in four phases, which are diagrammed in Figure 9. Throughout the item selection process, a Microsoft SQL Server database was used to manage the word and association lists, identify items for review, and monitor quality measures.

Phase 1 - Word list

A list of nouns was extracted from the Glasgow norms (Scott et al., 2018) with an age of acquisition rating of less than three and a familiarity rating greater than six (both rated on 7-point scales, see Table 6). These values were chosen through experimentation to minimise age of acquisition and maximise familiarity while still producing a variety of words. For reference, the average age of acquisition of PPTT items is 2.6 while the average familiarity rating is 5.9. Each noun on the list was then assigned a category from the list developed earlier (see Table 5).

Figure 9

Process of test item production



Phase 2 - Association list

A broad list of potential associations was generated from those nouns using a variety of methods including brainstorming pairs of list words, using a LSA 'Near Neighbours' calculation for each word, and brainstorming associations for each word from association types used in existing tests (e.g., used with, made from, grows into, same function). Consultation with colleagues occurred throughout this process to obtain a variety of ideas.

Phase 3 - Selected associations

A set of 71 test associations was then chosen from those generated. Selection of test items aimed to ensure the measured association, using LSA, was as high as possible, with an average association higher than that of the PPTT (0.34). Associations with a strength less than

0.3 were not considered, items with a strength higher than this were considered for selection. Selected associations also needed to be able to be depicted in a composite image.

Zieky (2015) discussed the importance of considering fairness throughout test design, rather than merely including minority groups in validation. As per the process used to select items for the short form of the PPTT (Breining et al., 2015), selected associations were based on defining features of the objects involved, rather than on cultural norms, stories, or stereotypes. This was intended to increase the likelihood of the test being able to be used across cultures. Zieky (2015) also presented a list of areas of knowledge likely to be unevenly distributed across groups (e.g., weapons, sports, religion) and practices likely to elicit distracting and irrelevant emotion in examinees (e.g., distressing topics, demeaning stereotypes). This list was reviewed and used to guide association selection. In particular, associations based on traditional gender roles, clothing norms, or stereotypes about cultural groups were not selected. Some examples of rejected associations are presented in Table 7.

Table 7

Examples of rejected associations

Rejected Association	Reason
Pool - Towel	Association too weak (0.19)
Bacon – Egg	Cultural food combination
Music – Piano	Not possible to depict abstract concept 'music'
Necklace – Woman	Gendered clothing norm
Ankle – leg	Not possible to depict alternatives
Ambulance – Hospital	Depiction requires cultural symbols
Arm – Gun	Distressing topic

Further, item selection aimed to balance the number of objects across categories, as in the Cambridge Semantic Battery (Adlam et al., 2010). Although the original intent was to select

an equal number of associations from each semantic category (listed in Table 5), insufficient potential items were generated to allow this. However, the number of associations from each semantic category was monitored throughout the selection process and more attention was given to finding associations from underrepresented categories.

After exhausting the initial list of associations (276 of which 209 were discarded as discussed), phases 1-3 were repeated including words with a familiarity rating between 5.8 and 6.0 and an age of acquisition rating between 3.0 and 3.2. Two further associations were selected during this process.

Phase 4 - Associations with distractors

Distractors were then generated for each association to produce the test items, selecting list nouns from the same category as each target object and considering visual similarity and how the associations could be depicted in a composite image. The strength of association with the target's associated object was again measured using the same LSA model as the previous step. Selection of distractors aimed to ensure the association with the target object was as low as possible and always lower than the target association. The new test aimed to have a larger average difference between the target and distractor associations than that of the PPTT (0.16).

Items where the association with either distractor was greater than 0.33 or the difference between the target association and either distractor association was less than 0.15 were individually reviewed. Following review, items were either accepted or new distractors were selected, and the process repeated. Items could be accepted if there were no viable

alternatives or if a strong distractor association arose from another meaning of a word which would be clearer in an image (e.g., a bridge crossing *over* a fire rather than currently *on* fire).

One teaching item was also generated, using an association for which the composite image is likely to be familiar to most people from weather forecasts (the sun behind a cloud) and distractors from different semantic categories (fridge and pencil).

Image Stimuli

Composite images were then needed for each test item. As in previous stages of test design, some specific quality considerations were identified and used to guide image selection.

1. **Clarity**. The images needed to be simple and clear. Examinees should be able to easily identify the key features of the objects depicted so that their answers reflect their semantic ability more than their visual acuity. To support this aim, it was decided to use black and white line drawings, rather than more realistic pictures or photos (Mazoyer et al., 1999).
2. **Consistency**. The image stimuli should not create confounding variables influencing examinees responses (Heuer & Hallowell, 2009). Differences in style, colour, size, line density, or positioning could suggest an answer. In addition to overall image consistency, the 'common' object should be identically depicted in all three answer options.
3. **Plausibility**. Each composite image should appear potentially correct. Thus, examinees are required to differentiate specific features in order to answer correctly, rather than simply choosing the option in which the two objects 'fit

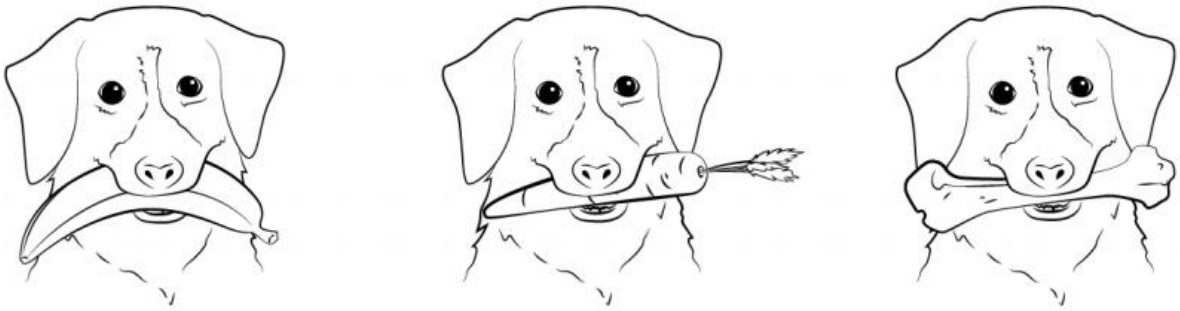
together' appropriately. For example, a saddle should appear to be positioned just as comfortably on a goat's back as on a horse.

Image resources online were examined to determine whether stock images could be acquired for the test. However, it was difficult to find large numbers of images in a sufficiently consistent style. Available licenses for purchased stock images were also examined. These generally restricted types of use and almost always included a clause reserving the right to revoke the license at any time. Due to these issues, it was decided to commission artwork specifically for the new test.

An artist was recruited through the Massey University Design School, using funding from the Massey Postgraduate Research Fund. An initial meeting with the artist was held to explain the rationale for the test design and image quality considerations. Regular meetings between the artist and the researcher continued throughout image production to discuss ideas and difficulties as they arose. As images were drawn, they were individually reviewed and discussed. Changes were then made as needed and the images reviewed again until accepted. Where there was uncertainty or difficulty, images were also reviewed and discussed with research colleagues. An example test item with final artwork is shown in Figure 10.

Figure 10

New test item with final artwork



Results

A total of 72 test items were developed, including the teaching item. The planned quality measures were then calculated for all PPTT and new test items and the results were compared.

The numbers of objects included in test items from each semantic category are depicted in Figure 11. One sample chi square tests were conducted for each item set to measure the evenness of spread and both item distributions were significantly uneven (p<.001). While more even category representation was desired, more potential items from under-represented categories could not be identified. Semantic categories are also unavoidably somewhat subjective and both item sets could be grouped in other ways, increasing or decreasing the measured evenness of spread.

Figure 11

Proportion of objects from each semantic category depicted in PPTT and new test items

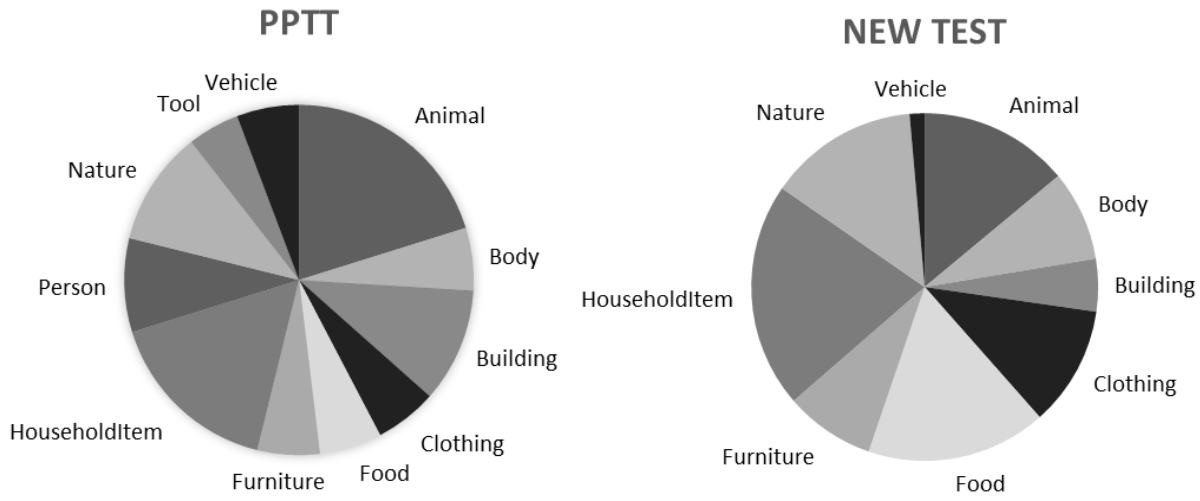


Table 8

Item quality measures of the PPTT and the new test items

Measure	PPTT mean (SD)	New test mean (SD)	
Word familiarity rating	5.9 (0.71)	6.4 (0.25)	***p*<.001**
Word age of acquisition rating	2.6 (0.81)	2.2 (0.37)	***p*<.001**
Associative strength of item targets	.33 (.15)	.45 (.08)	***p*<.001**
Associative strength of item distractors	.22 (.15)	.21 (.08)	*p*=.52
Difference between target and distractor associative strengths	.15 (.14)	.24 (.09)	***p*<.001**

As shown in Table 8, the mean familiarity ratings of the names of objects used in these items was significantly higher than the ratings of words used in the PPTT, while the mean age of acquisition rating was significantly lower (based on a two-tailed Students' t-Test). Further,

these PPTT figures exclude 20 words which were not present in the Glasgow psycholinguistic norms and are therefore likely to be even less familiar. Both measures suggest the new test items depict objects that will be familiar to most adults, regardless of their level of education.

Strengths of associations between objects in test items, measured using Latent Semantic Analysis (LSA), are also shown in Table 8. The mean associative strength of items in the new test was significantly higher than that of items in the PPTT, while the mean associative strength of distractors was similar, producing a significantly larger mean difference between target and distractor association scores. These measures suggest the new test items depicted well-accepted associations while the distractor objects chosen are not commonly related to the target objects. In contrast, distractors were more strongly associated than targets in nine out of the 52 PPTT items and a further six items had an association strength difference of less than .05.

The variability of items on all five metrics (as measured by the standard deviation) was noticeably lower for the new test than the PPTT, which suggests the new test items were more consistent and all of similar difficulty. Overall, the new test appeared to meet most design objectives and improve on the PPTT across a range of indicators of item quality.

Chapter 5: General Population Validation

This chapter discusses the methods and results of the first validation study for the new test. Because testing in awake craniotomy requires tasks that can normally be completed quickly and easily, validation in the general population aimed to determine whether the new test could be completed correctly by healthy adults from a variety of backgrounds. Key questions were:

- Do all test items perform similarly?
    - Intended outcome: yes
- Can most people score near ceiling on the new test?
    - Intended outcome: yes
- Do age, education, or culture influence scores on the new test?
    - Intended outcome: no

The methodology for recruitment and data collection will be described first, followed by the item performance results. These results were then used to inform a refinement of the item set, the process and outcomes of which will be discussed followed by overall test scores and group differences.

Ethics

A Massey University Human Ethics Low Risk Notification was completed for this part of the project (number 4000021031) (see Appendix B).

Survey Design

An online survey was selected as the simplest method to support recruitment of a large international sample. The survey was created using the Qualtrics platform and consisted of:

1. An information sheet providing information about the study, exclusion criteria, and a request for consent.
2. Five demographic questions: gender, age, ethnicity, country of residence, and highest educational qualification.
3. An instructional page for the test, including the sample item.
4. The 71 proposed test items.
5. An open request for comment on the test items and any difficulties experienced.
6. A thank you page, including a link to the study's Facebook page.

The full survey is reproduced in Appendix C. As well as responses to items, the survey also recorded the length of time participants spent on each test item. Both the order of the test items and the order of images within each test item were randomised for each participant.

Demographic questions

The wording of the ethnicity question and the ethnicity groupings chosen were based on the guidelines from the US Census Bureau (Matthews et al., 2017), modified to explicitly include New Zealand Māori. The New Zealand (Statistics New Zealand, 2021) and United Kingdom

(Government Statistical Service, 2015) guidelines were also reviewed. The US guidelines were chosen as the most suitable basis for an international survey because they included detailed rationale and evidence for the question design and also had the least country-specific guidelines. In contrast, the UK guidelines specify different ethnicity options depending on whether a survey is conducted in England, Scotland, Wales, or Northern Ireland.

The educational qualifications listed were based on the levels described by the International Standard Classification of Education 2011 (UNESCO Institute for Statistics, 2012). The wording of the items was designed to include words likely to be familiar to participants in both New Zealand and overseas.

A research colleague in the area of gender studies was consulted regarding the best wording of the gender question.

Recruitment and Participants

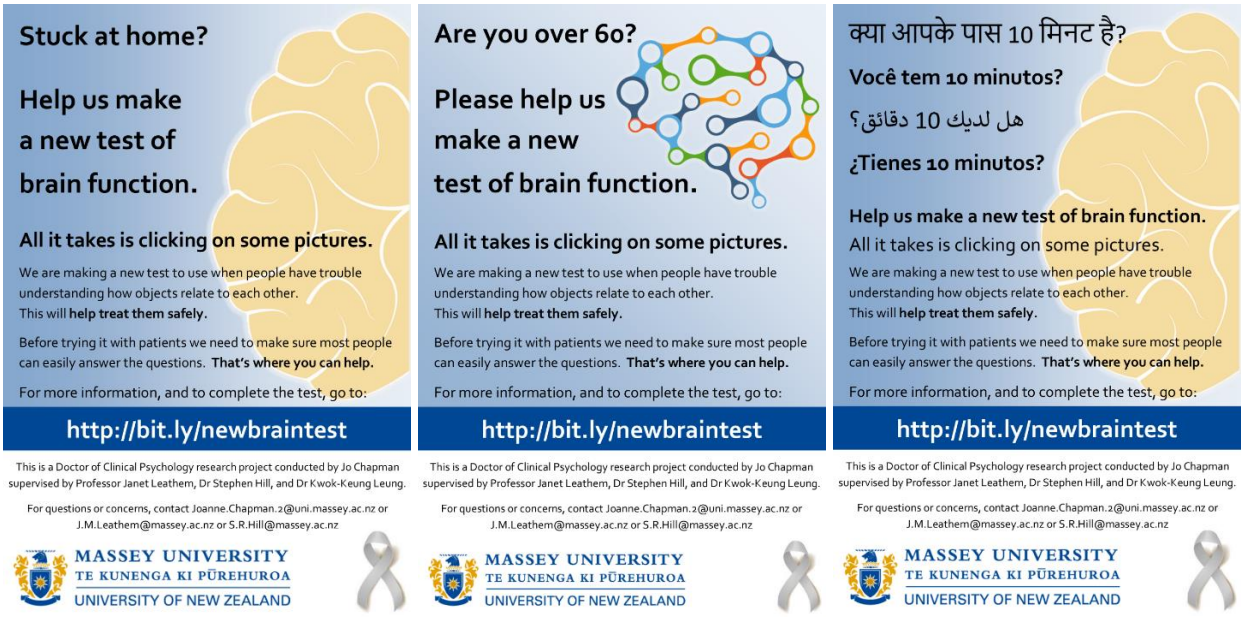
Recruitment aimed to obtain a large international sample of participants, with a broad range of ethnicities and all (adult) ages well represented. Participants aged 18 years and over were recruited online via snowball sampling, using social media and personal email contacts. A Facebook page was created for the study and was used to recruit participants, answer questions, and share information about the study's progress. Information about the study was also emailed to family members, friends, and selected community groups. The community groups contacted included GirlGuiding New Zealand's Trefoil Guild and six New Zealand cultural organisations.

Potential participants with experience of any significant neurological disease or injury were asked not to participate.

During data collection, the demographics of the sample to date were monitored and specific messaging was periodically used for more targeted recruitment of groups identified as under-represented. This included posts on the Facebook page and updated advertisements (see Figure 12). Older adults and people of minority ethnicities were both requested this way.

Figure 12

Participant recruitment advertisements



A total of 707 participants were included in the analysis: 700 who fully completed the survey and a further seven who completed at least 90% of the test items. While the sample was dominated by White/European women, all ages and education levels were well represented. The demographics of all 707 participants are summarised in Table 9.

Table 9

Demographics of study participants

Variable		n (%)
Gender		
	Male	148 (20.9)
	Female	549 (77.7)
	Other	8 (1.1)
Age		
	18-29	146 (20.7)
	30-29	168 (23.8)
	40-49	114 (16.1)
	50-59	72 (10.2)
	60-69	103 (14.6)
	70-79	71 (10.0)
	80+	31 (4.4)
Education		
	No formal qualification	21 (3.0)
	Some secondary education	76 (10.7)
	Completed secondary education	92 (13.0)
	Trade / Vocational qualification	86 (12.2)
	Undergraduate degree	263 (37.2)
	Postgraduate qualification	164 (23.2)
Ethnicity*		
	White / European	622 (88.0)
	Māori / Pasifika	45 (6.4)
	East Asian	35 (5.0)
	Indian / South East Asian	23 (3.3)
	Other	28 (4.0)
Country of residence	New Zealand	543 (78.7)
	Australia	44 (6.4)
	United Kingdom	30 (4.4)
	United States of America	29 (4.2)
	Hong Kong	15 (2.2)
	Other	29 (4.2)

* participants could endorse multiple ethnicities

Inspecting these demographics for interactions revealed that participants aged over 80 were more likely to have no formal education, those aged 18-29 more likely to have completed

secondary education, and those aged 40-49 or of Asian ethnicity more likely to have postgraduate qualifications. Finally, there were more White/European participants aged 70-79 and fewer aged 30-39 than in other age groups.

Ethnicity coding

Several participants gave their nationality as an 'Other' ethnicity (e.g., "New Zealander", "Scottish", "Chinese") and these were recoded into the relevant groups before analysis (e.g., East Asian, White/European). Further, while best practice ethnicity data collection allows participants to endorse multiple ethnicities, the statistical analyses used in this study required each participant to be assigned to a single ethnic group. Participants were not asked to identify a 'main' ethnicity and this sample was deemed too small to analyse the subgroups formed by combinations of ethnicities (e.g., European-Māori and European-Asian). Therefore, a simplified prioritisation approach was used to produce the required data structure while maximising the visibility of minority ethnicities in the analysis (Cormack et al., 2011). Participants who endorsed White/European alongside any single other ethnicity were coded as their non-white ethnicity. Participants who endorsed two or more non-white ethnicities were coded as Other. Finally, ethnicity groups with too few participants for analysis (e.g., African, Hispanic) were grouped into 'Other'. Some examples of recoded ethnicity responses are shown in Table 10.

Table 10

Examples of recoded ethnicity responses

Original response(s)	Final ethnicity code
"Pākehā", "New Zealander"	White/European
"Scottish", "Dutch"	White/European
"Chinese"	East Asian
White/European AND Māori/Pasifika	Māori/Pasifika
Māori/Pasifika AND East Asian	Other
African	Other

Analysis of Item Performance

Almost half of participants (339) wrote a comment on their experience of the test items. Comments made by participants were each read and coded using a simple content analysis approach. Most codes reflected difficulties with specific items and a frequency count of these codes was calculated.

Item analysis then aimed to identify any items performing differently to the test as a whole. Three performance indicators were used for each item:

- Count of incorrect or skipped responses
- Median time spent in seconds
- Count of participant comments mentioning difficulty with the item

All three measures in this analysis had a wide range of values. Modified z-scores were used due to their ability to describe variance in samples with high variability without being overly affected by large outliers (Leys et al., 2013). These z-scores are based on the median and median absolute difference (MAD), rather than the mean and standard deviation of a sample. In addition to their high variability, the data were also asymmetric with a far wider range of

values above the medians than below. To account for this, a modification of the 'Double-MAD' approach (Rosenmai, 2013) was used to calculate a MAD for values above the median only. As an illustration of the effect of this approach for this sample, these statistics are compared in Table 11. All calculations were conducted using Microsoft Excel.

Table 11

Comparison of descriptive statistics using this study's data

	Incorrect/Skipped answers	Median time spent (seconds)	Participant comments
Lowest value	0	3.0	0
Highest value	127	10.3	52
Mean (SD)	12.06 (23.82)	4.5 (1.4)	5.1 (10.2)
Median (MAD)	3 (3.0)	4.0 (0.6)	1.0 (1.5)
Median (MAD above median only)	3 (6.8)	4.0 (1.0)	1.0 (3.4)

Modified z-scores were calculated for each measure for each item. For the time spent measure, the median of median times spent was used as the basis. The z-scores were then graphed so that patterns could be observed and items scoring highly could be identified and examined. This graph is shown in Figure 13. Items with a z-score greater than 2.0 (two MAD above the median) on any measure were flagged for analysis. Two MAD above the median was selected as a conservative threshold, following Leys et al.'s (2013) recommendation. This resulted in a list of 17 flagged items.

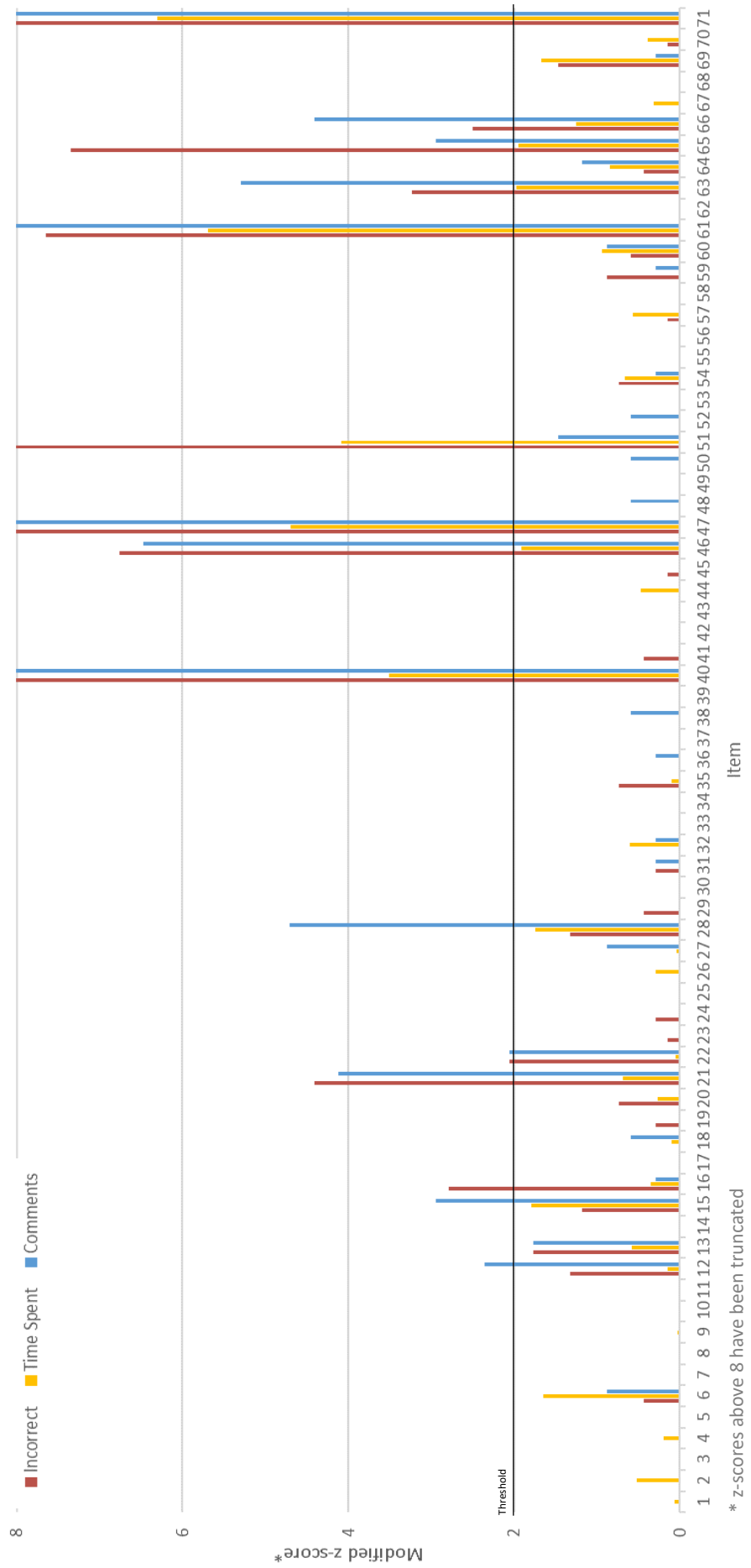
The responses to each item were also examined to determine if any were disproportionately answered incorrectly (or skipped) by people of different ages or ethnicities. The percentage of participants of each ethnicity/age answering each question incorrectly was

calculated and the results inspected for uneven patterns. Seventeen items did not appear to perform consistently between groups.

Combining the lists of items identified by these two analyses (diagrammed in Figure 15) resulted in 22 items for individual examination. The remaining 49 items were accepted as working as designed.

Figure 13

Modified z-scores for incorrect/skipped answers, time spent, and comments made on each test item



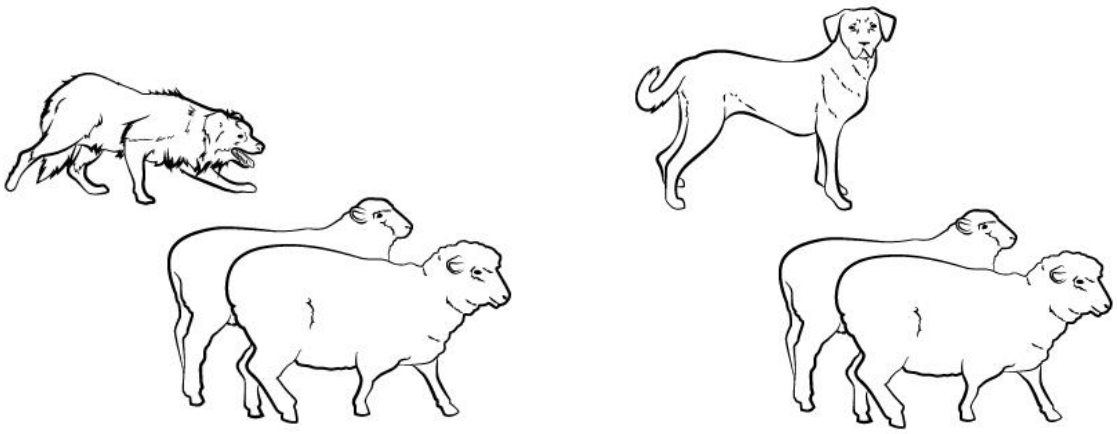
Item Refinement

The 22 items flagged by the item performance analysis were each examined in turn. Possible causes for problems were theorised, informed by distractors selected as incorrect responses, close reading of participants' comments, and discussions with colleagues. Consideration was then given to whether the issues with each item were acceptable, could be solved, or necessitated removal from the test.

Most flagged items were removed as solutions either could not be identified or required such a large change that re-validation would be necessary (e.g., changing a distractor from a cat to a cow). Six items were retained with the image stimuli updated to improve clarity (e.g., changing the breed of dog, increasing the relative size of image components). One of the updated images is shown in Figure 14. Three items were retained unchanged. Two of these appeared to have small cultural effects but no cause could be theorised, so were regarded as possible measurement error. The final item was frequently commented on by participants as there were two possible interpretations of the images (kissing a cheek or whispering in an ear), however, both interpretations led to the correct answer. A Trello board was used to record the issues identified and decision made for each test item. Appendix D details the specific issues, comment themes, and decision rationale for each of the 71 proposed test items. At the end of this process, diagrammed in Figure 15, 58 items were included in the final test.

Figure 14

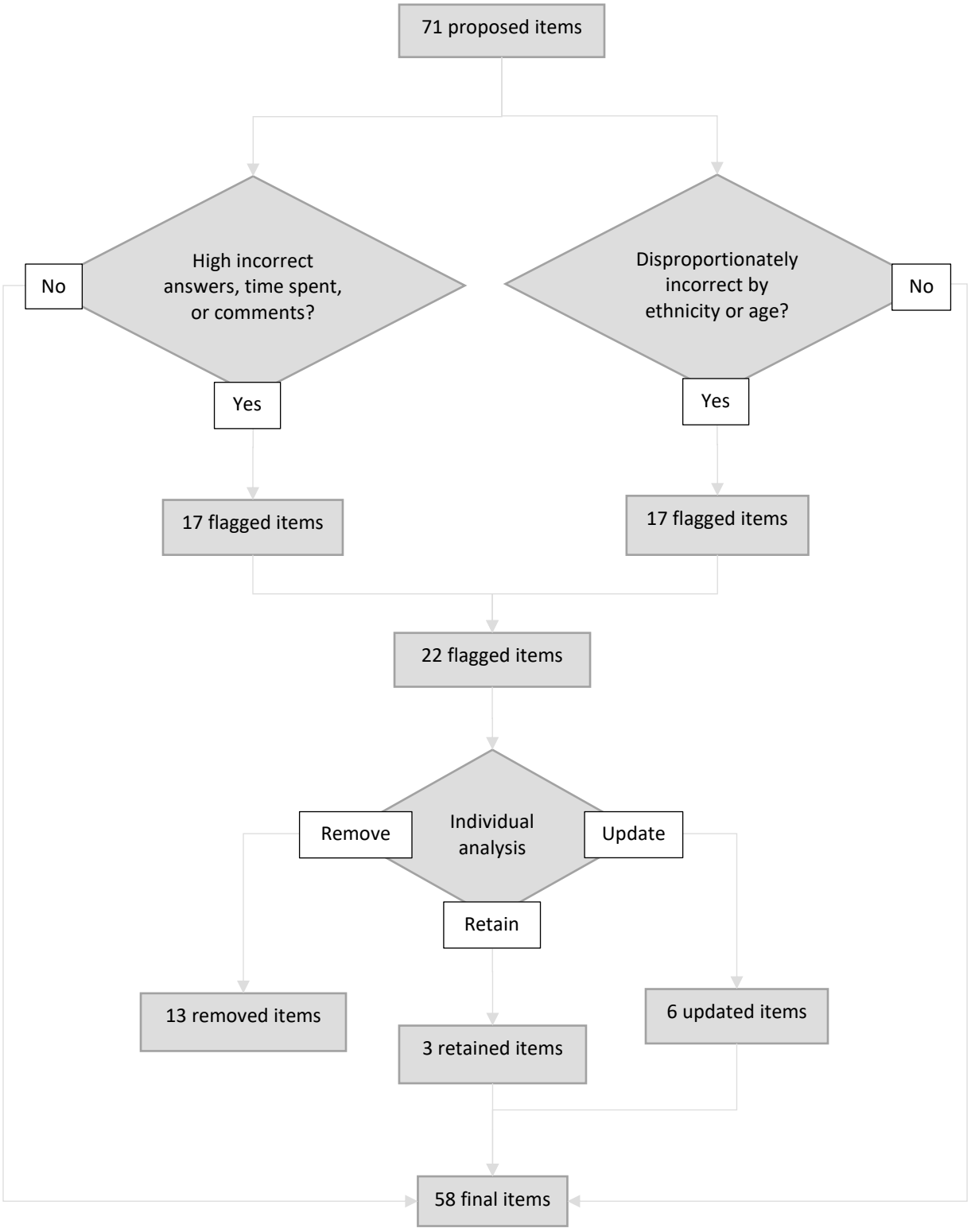
Example of an image updated after item performance analysis



Original Updated

Figure 15

The process of analysis used to produce the final set of 58 test items



Updated test quality measures

The test quality measures used in the previous chapter were recalculated for the reduced 58-item set and again compared with the PPTT. While more food-related items were removed than other categories (see Figure 16), the significant unevenness of category representation across the test items was broadly unchanged. Average familiarity, age of acquisition, and associative strengths were nearly identical to the original item set, with significant differences from the PPTT preserved (using a two-tailed Students' t-Test; Table 12).

Figure 16

Proportion of objects from each semantic category depicted in PPTT and new test items

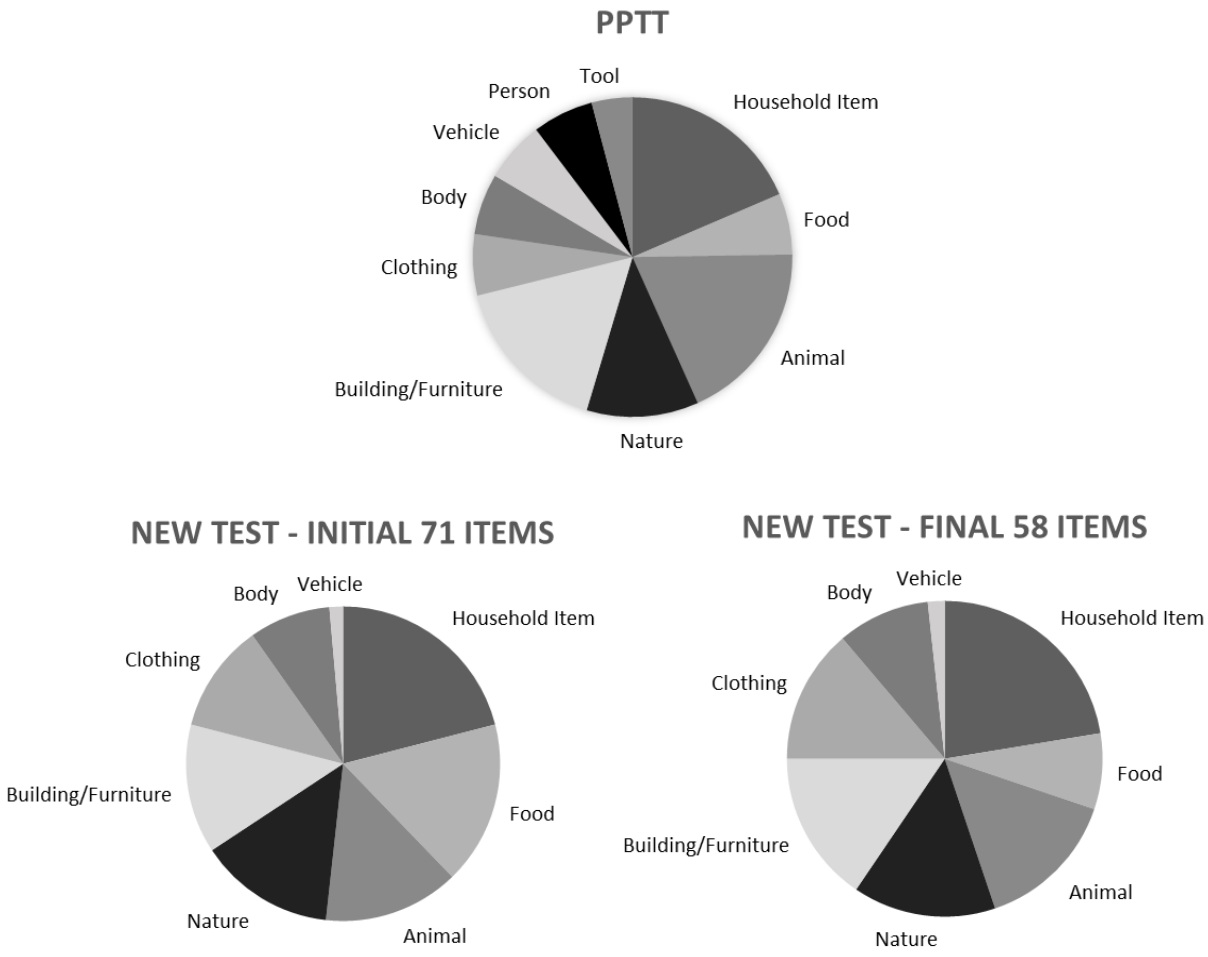


Table 12

Item quality measures of the initial and final new test items, compared with the PPTT

Measure	New test		PPTT	
	Initial 71 items mean (SD)	Final 58 items mean (SD)	mean (SD)	
Word familiarity rating*	6.4 (0.25)	6.4 (0.25)	5.9 (0.71)	***p*<.001**
Word age of acquisition rating*	2.2 (0.37)	2.2 (0.36)	2.6 (0.81)	***p*<.001**
Associative strength of item targets	.45 (.08)	.46 (0.09)	.33 (.15)	***p*<.001**
Associative strength of item distractors	.21 (.08)	.21 (0.08)	.22 (.15)	*p*=.55
Difference between target and distractor associative strengths	.24 (.09)	.24 (0.09)	.15 (.14)	***p*<.001**

*Rated on a scale from 1-7, see Chapter 4 for details

Overall Test Performance and Group Differences

All participant's tests were rescored to exclude the items removed in the previous stage and further analysis was conducted using only the final 58-item set. The median score was 58/58 with 78.9% of participants answering all items correctly and a further 16% of participants making only one error. This suggests that, as intended, most healthy adults could easily complete the test.

Many statistical tests were inappropriate for group comparisons due to this ceiling effect and uneven distribution (Šimkovic & Träuble, 2019). Pearson's chi-squared test can validly interpret not only skewed data but also subgroups with small sample sizes (McHugh, 2013). Analyses were conducted using IBM SPSS Statistics, dividing participants into two groups: those answering all questions correctly and those who gave one or more incorrect

responses. Finer-grained divisions of scores were not possible due to the very small number of participants who made more than one error. Where necessary, small demographic subgroups were combined until a valid chi-squared statistic was able to be calculated. All chi-square analyses were later recalculated using a two-tailed Fishers Exact Test, with the same findings in terms of significance.

No significant group differences were found for gender or level of education. There were also no significant differences between White/European participants living in different countries. However, significant chi-squared test results ($p<.05$) were obtained for the ethnicity and age variables, suggesting that age and culture may affect test scores.

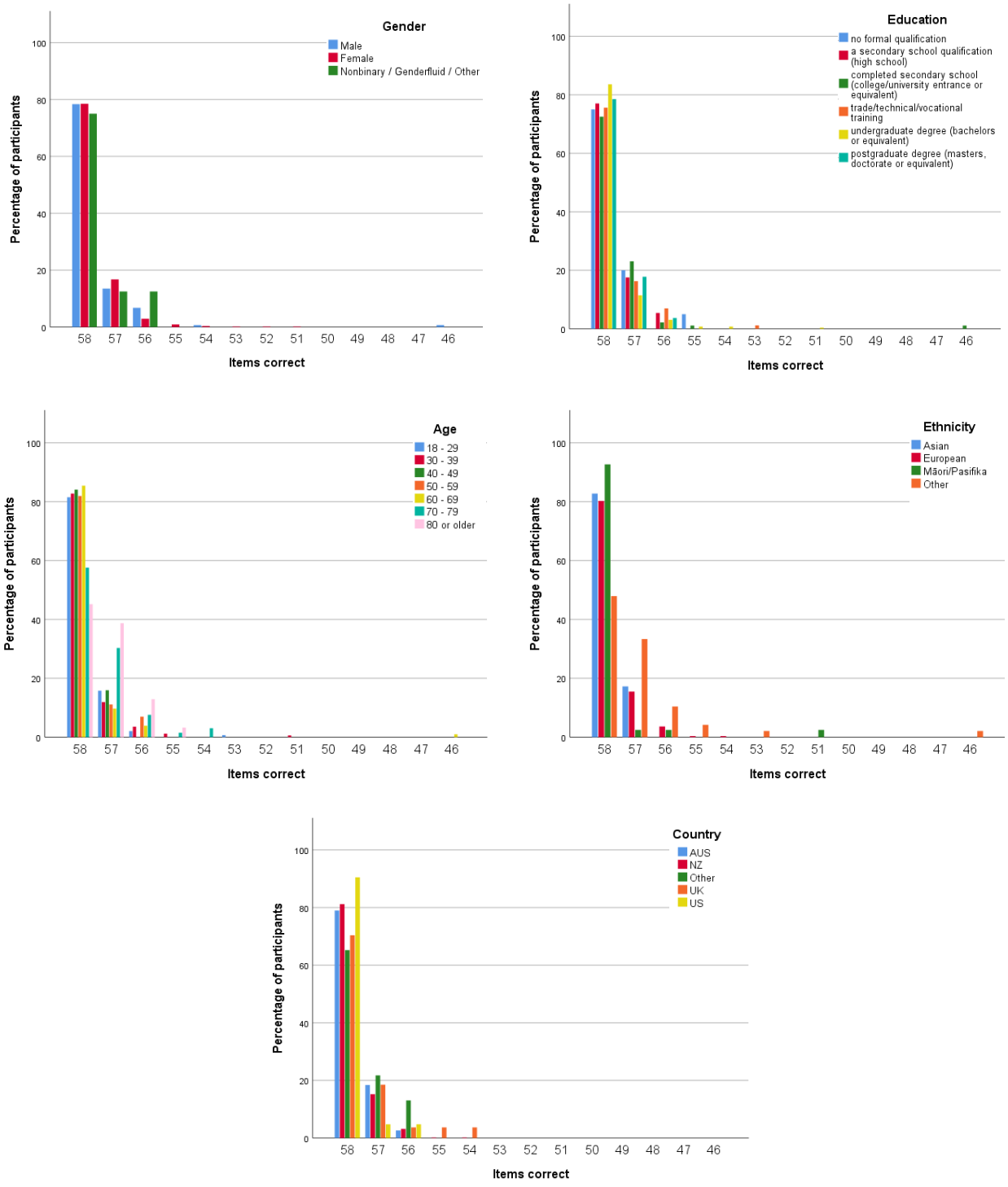
Given the relatively high numbers of cells in the two significant tables (2x4 and 2x7), the post-hoc analyses described by Sharpe (2015) were used to explore potential sources of the significant results. For ethnicity, standardised and adjusted residuals identified the White/European and Other groups as those with the greatest discrepancies. A Bonferroni adjusted z-test was also significant for cells involving these groups. Ransacking, by dividing the data into smaller sub-tables of interest, returned a significant result when comparing the European and Other groups with each other, but found no significant differences between the European, Asian, and Māori/Pasifika groups. These results are congruent with a visual inspection of the data (see Figure 17) and suggest that participants in the 'Other' ethnicity group were significantly less likely to answer all items correctly, while the remaining ethnic groups scored similarly. Repeating this analysis process for the age variable suggested participants aged 70-79 were significantly less likely to answer all questions correctly. Visual inspection (see Figure 17) and the Fisher's Exact Test results suggested this may also be the

case for participants aged over 80, although neither Chi Square cell residuals nor z-tests supported this conclusion.

In summary, the new test was generally completed accurately by participants of: all genders; all educational backgrounds; European, Māori/Pasifika, or Asian ethnicities; and ages 18-69. Participants from other ethnic groups and/or aged over 70 were more likely to receive lower scores. However, as shown in Figure 17, many of these 'low-scoring' participants gave only one incorrect response. It was concluded that, while some statistically significant age and culture effects were present in the new test, these effects were so small that they may be clinically insignificant. In contrast, age and culture effects of the PPTT have been documented to produce incorrect responses to up to 23% of its items (Gamboz et al., 2009)

Figure 17

Test scores by gender, education, age, ethnicity, and country of residence



Chapter 6: Clinical Validation

This chapter discusses the planned method and actual outcome of the second validation study for the new test. Validation in clinical populations aimed to determine whether the new test was a valid measure of semantic and/or IFOF impairment. Mapping critical brain areas in awake craniotomy requires any errors to be valid indicators of stimulation-induced impairment. Two relevant clinical groups were identified: glioma and specific semantic impairment.

Individuals with gliomas were the target group for the use of the new test so were also appropriate for its validation. The firm aim was to investigate convergent validity by comparing their scores on the new test with their scores on existing semantic tests. Further, if participants with glioma had more difficulty with the test than the general population, that could indicate the test was sensitive to glioma or treatment-related cognitive changes. Finally, if those whose IFOF was adversely affected by their tumour had most difficulty, this would provide evidence that the test was sensitive to IFOF impairment. However, many tumours do not produce cognitive symptoms. Patients may have had little difficulty with the test, even if it was working as designed, so trials with this group were not sufficient to examine its validity.

Thus, validation was also planned with individuals who had specific semantic impairments. These participants may have experienced small strokes in key areas of the semantic network or have a diagnosis of the semantic variant of Primary Progressive Aphasia (svPPA). People with svPPA have been the focus of the majority of studies of semantic memory (Patterson et al., 2007; Rogers, Lambon Ralph, Hodges, et al., 2004) and could be expected to score significantly lower than healthy adults on a valid test of non-verbal semantic association. The same hypothesis could be made concerning individuals with specific semantic impairments after localised strokes (Han et al., 2013).

Participants from both groups were asked to complete both the new test and two existing semantic tests. Key questions were:

- Are scores on the new test related to scores on existing semantic tests?
    - Intended outcome: yes, comparable or positively related scores
- Do people with semantic impairment score lower than the general population?
    - Intended outcome: yes
- Do participants with glioma score lower than the general population?
    - Either outcome possible
- Does IFOF tumour involvement influence scores on the new test?
    - Intended outcome: yes, IFOF involvement related to lower test scores

The processes for ethical review, recruitment, and data collection will be described first, followed by the results.

Ethical Review

This study was reviewed and approved by the New Zealand Health and Disability Ethics Committee (HDEC) (see Appendix E). The determination of the appropriate ethics committee for the study was informed by a Scope of Review submission to HDEC. A key ethical consideration was the engagement with potential participants whose neurological conditions may have diminished their capacity to consent to the research. Other particularly relevant ethical principles were privacy/confidentiality, just distribution of burden, and informed consent without undue influence. Discussions with the committee further explored the design of the consent process as well as the protection of the researcher's intellectual property in the new test, the validity of neuropsychological testing via video conference, managing participant distress, and validating the test with Māori. Informed consent was the most complex of the issues addressed and is described further below.

Informed consent

Consideration was given to preventing the exploitation of vulnerable patients while also avoiding unnecessary denial of patients' rights to participate in research. Discussions with supervisors, experienced Clinical Psychologists, a Speech Language Therapist, and Massey University Research Ethics staff informed the researcher's thinking on these issues. It was concluded that most potential participants would be capable of deciding whether to participate in this research with an appropriately supported and accessible process (the final process used is discussed in detail below). The types of tasks involved in the study (completing tests, sharing medical information) were likely to be familiar to potential participants and they may have consented to similar activities often during their medical treatment.

On the advice of the consulted Speech Language Therapist, a version of the participant information sheet was created in EasyRead format. This format is designed to be more easily understood by people who have difficulty with typical written text (People First New Zealand, 2014). Images for the EasyRead information sheet were acquired from the "easy on the i" image bank (Leeds and York Partnership, 2021).

Planning a supported and accessible consent process was complicated by the need to conduct this study online due to the Covid-19 pandemic. Tools and guidelines for telepsychology and digital consent were reviewed (Agency for Clinical Innovation, 2019; American Psychological Association, 2020; Bilder et al., 2020; Joint Task Force for the Development of Telepsychology Guidelines for Psychologists, 2013). Most tools reviewed required participants to create accounts, install software, and/or set up identity verification before meeting with the researcher. Reducing cognitive burden placed on participants was a priority, both to accommodate potential impairments (likely to affect patients' ability to remember and follow through with tasks) and to respect their time/energy during what was likely to be a very stressful time. Thus, complex software tools and signed consent forms mailed to the researcher after the meeting were all considered inappropriate for this study. However, robust evidence of consent and identity was still needed, especially given participants' potential vulnerability and the need to request medical information. The procedure described by Welch et al. (2016) provided a potential solution and was used as the basis for this study's consent process. There were two key components: a digital consent form, completed while meeting with the researcher, and a photograph of the participant completing the form as additional evidence of their identity.

The final consent process used for this study is diagrammed in Figure 18, while Table 13 lists relevant ethical considerations and standards with their resulting process components.

Figure 18

Consent process

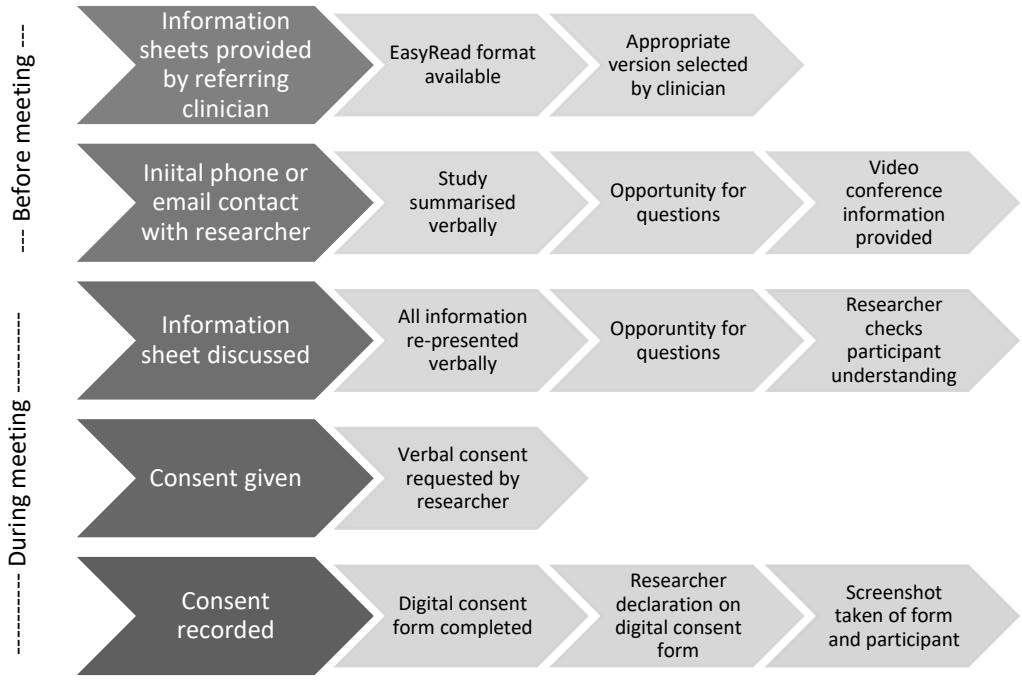


Table 13

Ethical considerations/standards and related consent process components

Consideration/Standard	Consent process components
Supporting participants' potentially reduced capacity to consent	Information provided in multiple forms (written, EasyRead, oral, and video). Participants encouraged to have a support person present. Questions asked of participants to check their understanding throughout the consent process.
Consent without undue influence	Consent process entirely completed by the researcher, who was not involved in participants' medical care.
Reducing burden on participants by accommodating likely cognitive difficulties	Consent process entirely completed with the researcher present to assist. Participants *not* asked to remember/organise to complete tasks after the meeting or to use unfamiliar/complex digital identity software.
Not recording the participant without their consent	Verbal consent *not* recorded. Image of participant only recorded at the end of the consent process.
"Electronic consent methods must entail a means to ensure that the participant himself or herself provided consent." (National Ethics Advisory Committee, 2019, p. 79)	Digital consent form includes a field in which the participant can draw a version of their signature. Image recorded of participant.
"When utilising electronic consent, interaction between researcher and participant should remain an integral part of the consent process." (National Ethics Advisory Committee, 2019, p. 79)	Consent process entirely completed during a meeting with the researcher.
"…a permanent record of the process, as evidence that information was provided in an appropriate manner and informed consent was obtained free from coercion" (National Ethics Advisory Committee, 2019, p. 61)	Digital consent form saved. Digital consent form includes a researcher declaration.

Participant Selection and Recruitment

Participants were recruited from two groups: patients with current or previous glioma, and patients with a diagnosed semantic impairment. To help determine whether the new test was specific to semantic difficulties, potential participants from either group were ineligible to participate if they had other conditions which could affect their test performance, specifically:

- any *other* significant language impairment
- a visual or perceptual impairment (not corrected by glasses)
- any *other* significant neurological condition (e.g., traumatic brain injury, dementia)

Referring clinicians discussed the study with their patients during regular appointments or ward rounds and provided appropriate information sheets (see Appendix F). If patients expressed interest in participating, the referring clinician forwarded their contact details to the researcher. During first contact with the researcher (by either email or phone), patients were again given a summary of the study and had the opportunity to ask questions. If they were still interested in participating, the researcher arranged to meet with them via video conference at a time which was convenient. Instructions for using video conferencing were provided if needed and the researcher was available to assist with any technical problems.

Recruitment through Wellington Regional Hospital

Relationships were developed between the research team and the Clinical Lead Neurosurgery from Wellington Regional Hospital. Study recruitment materials and information about eligibility criteria for both groups (see Appendix F) were provided and discussed with Department of Neurology department staff during a regular training session. Doctors (primarily

the Neurosurgeon and his Registrars but also Neurologists) identified patients eligible for the study.

A further relationship was developed with the Wellington Regional Hospital Team Leader for Speech-Language Therapy and Dietetics and information was provided through her to all Speech Language Therapists (SLTs) at the hospital. At her invitation, a presentation about the study was also made to the Aphasia Special Interest Group, consisting of SLTs from the Wairarapa, Hutt Valley, and Capital & Coast District Health Boards.

Other recruitment

Advertising material (see Figure 19) and information about the study (see Appendix F) was distributed by the New Zealand Speech Language Therapists' Association through their Facebook page and regular newsletter. An SLT known to the researcher also shared the advertisement with their professional colleagues and Aphasia New Zealand agreed to share information about the study with their members, both affected families and clinicians. Finally, the researcher presented the study at the 2021 scientific meeting of the Neurological Association of New Zealand, inviting attendees to refer eligible patients (see Appendix H).

Figure 19

Advertisement for Speech Language Therapists

Sample size

Recruitment for this study was expected to be challenging as glioma, svPPA, and other specific semantic impairments are all relatively rare conditions and New Zealand's population is small. Given the incidence of glioma in New Zealand and the population served by Wellington Regional Hospital (around 900,000), an estimated 37 people could present with glioma each year. When this project began, the hospital conducted 5-10 awake craniotomy surgeries annually. The potential population with specific semantic impairment was more difficult to estimate as it could be caused by many different conditions. Data on the incidence of svPPA is limited and estimates vary widely from 1.25 up to 2.81 cases per 100,000 person-years (Coyle-

Gilchrist et al., 2016; Hogan et al., 2016). These estimates suggest 11-28 potential cases in the Wellington Regional Hospital area and 64-144 in New Zealand as a whole. Estimates of the proportion of stroke survivors experiencing specific semantic impairment could not be found.

With Wellington Regional Hospital as the recruitment site for glioma patients and recruitment for the semantic impairment group occurring New Zealand wide, we felt a total sample of at least 10 participants could be achievable.

Study Procedure

Participants met with the researcher once, via Zoom video conferencing, and were encouraged to have a support person present. The meeting agenda was as follows:

1. *Discuss the participant information sheet.* This discussion aimed to support and check the patient's understanding throughout. Both they and their support person were encouraged to ask questions of the researcher.
2. *Request and record consent to participate.* Once given the participant's verbal assent, the researcher displayed an online consent form. The participant was given control of the screen to complete the form with the support of the researcher. When the form had been completed, the researcher saved a 'screenshot' as evidence of the identity of each participant.
3. *Collect demographic information.* A second online form was displayed and participants were asked to indicate their gender, age, ethnicity, and highest educational qualification. The wording and answer options for these questions

was the same as that used in the general population study (see Chapter 5 for a detailed discussion).

4. *Complete neuropsychological tests.* The researcher administered three tests, the new test and two comparison tests, which are discussed further below.
5. *Debrief and thanks.* Participants were offered an opportunity to comment on their experience of the tests and/or ask questions of the researcher. This allowed a debrief of any distress arising from finding the tests difficult.

A printed document, with administration instructions and reminders of key steps, was used to guide the researcher through the agenda and ensure a consistent process for each participant. This document is reproduced in Appendix G. The consent form, demographics form, and test stimuli were implemented using Qualtrics surveys. The final survey also included spaces for the researcher to record comments made by participants or observations during test administration. These questions were not displayed to the participant and were completed after they had left the meeting. Screenshots of surveys are included in Appendix G.

Following the meeting, the researcher emailed the participant's referring doctor or SLT requesting specific medical information and providing evidence of their consent to this information being shared. The following information was requested:

1. For participants with glioma:
    a. The general location of the tumour (e.g., left superior frontal lobe).
    b. The grade of the tumour.
    c. Time since onset of the tumour.

d. Whether, based on medical imaging, the participant's IFOF was invaded, displaced, or spared by the tumour.

e. Whether/When the participant had had surgery or radiation treatment.

2. For participants with semantic impairment:

a. Diagnosis (only that causing the semantic impairment, not any other diagnoses).

Test Administration

All participants completed three tests: the new test, the Pyramids and Palm Trees Test (PPTT; Howard & Patterson, 1992), and verbal fluency (using F, A, S, and Animals). The PPTT was selected as the test intended to be replaced by the new test while verbal fluency was used as a more valid comparison measure of semantic impairment. See Chapter 2 for a detailed discussion of the PPTT and verbal fluency.

The researcher used standardised administration protocols for all three tests (see Appendix G) and stimuli were displayed using screen sharing. The order of the two non-verbal tests was randomised, participants completed either the PPTT or the new test, then verbal fluency, then the other non-verbal test (PPTT or new test). Participants were not told which was the new test until after all tests had been completed. They were offered the choice of their support person either leaving the room during the tests or sitting behind the participant out of their line of sight. This prevented inadvertent signalling of answers or distraction by the support person.

Testing via video conference

The new test was administered on screen with the participant given control of the screen to select their answers or, where this was not practical, indicating their choice verbally by saying “left”, “right” or “middle”. This administration functioned in the same way as the online administration used in the general population study (see Chapter 5) except that the researcher was available to answer questions and encourage the participant. The PPTT was administered on screen in the same way as the new test while verbal fluency did not require any alteration to standard in-person administration. Video conference administration of visual tests has been associated with only a very small score decrease and has no effect on verbal fluency scores (Brearly et al., 2017). Video conference testing has also been shown to validly identify cognitive impairment (Wadsworth et al., 2018).

Updates to the new test

This study used the reduced item set and updated images resulting from the general population validation of the new test (see Chapter 5). That study randomised the order of items and stimuli for each participant but, for this study, a single standardised order was created using the List Randomizer and Integer Generator tools provided by RANDOM.ORG (2021). The test instructions were expanded for verbal administration, including responses to likely questions. These updates were intended to produce a standardised version of the test, ready for wider use (provided in Appendix A).

Planned analyses

As the test was intended to be extremely easy for healthy adults to complete, many statistical tests would be inappropriate for group comparisons due to ceiling effects and uneven distribution (Šimkovic & Träuble, 2019). Chi square analyses were originally planned, using a cut-off score of 1 or more errors (derived from the general population study) to divide participants into high and low scoring groups. As well as being able to interpret highly skewed data, chi square analyses can also be validly performed with a low sample size (McHugh, 2013). Later, Fishers Exact Test was considered as an alternative that would permit the interpretation of even smaller sample sizes (Singh, 2014). Based on the proportion of high and low scores found in the general population study, at least 5 participants were sought in each group to ensure no cell had an expected value less than 1.

Specific analyses planned were:

- New test scores compared with PPTT test scores
- New test scores compared with verbal fluency test scores
- New test scores of participants with semantic impairment compared with the general population
- New test scores of glioma patients compared with the general population
- If possible, new test scores of glioma patients depending on whether their IFOF was invaded, displaced, or spared by the tumour

Further analyses were considered if permitted by a sample size larger than expected. Scores could be correlated with by the percentage of the IFOF invaded by the tumour (as in Almairac et al., (2015)) and/or subgroup analyses could be performed by high/low grade of

tumour, time since tumour onset, and left/right hemisphere location. These analyses would aim to find a negative correlation between IFOF invasion and test scores for tumours in either hemisphere. It was also hypothesised that high grade tumours with more recent onset may have a greater effect on performance as the brain has had less time to adapt.

Results

Despite the range of recruitment efforts, only three participants were recruited over the study period of March-December 2021 (during the Covid-19 pandemic). All three participants had gliomas and were referred by the Neurology team at Wellington Regional Hospital. One referral from the Neurology team was declined as the patient was under 18 years of age and a fifth referred patient was too unwell after surgery to participate. One further referral from an SLT at Wellington Regional Hospital was transferred to another hospital before the study could be discussed with them. Information was forwarded to the patient's new hospital but they did not make contact. Two referrals from a Neuropsychologist colleague were declined as they did not meet criteria for either the glioma or semantic impairment groups. No referrals were received from other SLTs in the Aphasia Special Interest Group, the New Zealand Speech Language Therapists' Association, Aphasia New Zealand, or the Neurological Association of New Zealand. Potential contributing factors to the low referral rate are discussed in the following chapter.

Participants and the ethics committee were assured no individual data would be published, but some general statements can be made. All participants were from the glioma group and none had significantly more difficulty with category fluency than letter fluency,

suggesting they did not have semantic impairments at the time of testing. They engaged easily with the stimuli of the new test, were able to understand the task, and participated with little difficulty in administration via video conference. Error rates on the PPTT and new test appeared similar on visual inspection and were mostly within the normal ranges for both tests. With such a small sample size, no further analyses were possible and no conclusions can be drawn regarding the validity, specificity, or sensitivity of the new test.

Chapter 7: Discussion

This project aimed to develop and validate a new test of non-verbal semantic association, specifically for use during awake craniotomy surgery. Wellington Regional Hospital approached Massey University with a need for a replacement for the Pyramids and Palm Trees Test (PPTT), which was not performing adequately for their needs. During surgery, critical brain regions are 'mapped' by repeatedly administering tasks that can normally be completed quickly and easily. If the patient is suddenly unable to respond correctly when a single brain region is deactivated by stimulation, then that region is critical to performing the task and must be preserved. Different tasks are needed for different areas of the brain (based on their typical functions) and the PPTT was used to map the Inferior Fronto-Occipital Fasciculus (IFOF). Studies of the PPTT show that it has a range of problems, especially in its performance across cultures, and no other existing test met all requirements for effective use in IFOF mapping.

The most well-researched of the IFOF's functions is semantic memory, or 'general knowledge' of objects and concepts. Many theoretical models of semantic memory have been proposed, all of which emphasise the importance of wide-ranging connectivity between brain regions to integrate semantic knowledge and control its retrieval. The new test was intended to assess the ability to retrieve and integrate semantic information by evaluating associations

between objects. A novel test format, composite images, was conceived to ensure the new test was simple to understand and administer while also being complex enough for potentially high sensitivity and specificity to IFOF impairment.

Validation of the new test in the general population aimed to determine whether all items performed similarly and whether it could be easily completed by healthy adults from a range of backgrounds. A second study then aimed to investigate the sensitivity of the test to IFOF or semantic impairment in clinical populations, alongside comparing its performance with existing semantic tests.

Outcomes

Development and general population validation

Using preselected measures of item quality, the final test metrics were an improvement upon the PPTT in all areas and met the design aims. A total of 707 healthy adult participants were then recruited to complete the new test via an online survey. The 71 initially proposed test items were reduced to a final set of 58 items which performed consistently as intended. Analyses of subgroup performance found there were no significant differences between participants of: all genders; all educational backgrounds; European, Māori/Pasifika, or Asian ethnicities; and ages 18-69. Participants from other ethnic groups or aged over 70 were less likely to answer all items correctly, but generally made only one error. These results suggested that, as designed, most healthy adults could quickly and easily complete the test.

The robustness of processes for both item selection and norming were a strength of the development of the new test. In comparison to other semantic tests developed during the

same period (Cejudo et al., 2022; Luzzatti et al., 2020; Mazoué et al., 2022; Pozueta et al., 2019; Rey, 2020; Verst et al., 2021; Zannino et al., 2021), this project had a far larger normative sample and wider range of quality measures. No other test attempted to assess the cultural fairness of the items and less than half investigated age or education effects on test scores. Few measured familiarity or age of acquisition of the objects depicted and most did not attempt to quantify the relative strength of semantic associations selected.

The general population validation was limited by the predominantly New Zealand-based sample. While no clinically significant ethnicity effects were found, the test's potential performance in other countries is not certain. With regard to the development of the test itself, the bespoke black-and-white images created were intended to support clarity and consistency but more realistic images may have been quicker and easier for examinees to recognise (Cejudo et al., 2022; Heuer, 2016). Additional approaches used by other recent test developers in this area could have increased the potential sensitivity and specificity of the new test, notably selecting items based on their discriminating power (Pozueta et al., 2019) or providing item sets at multiple levels of complexity (Verst et al., 2021). Finally, Latent Semantic Analysis is far from the only way to measure semantic association and improving methods for computational semantic processing is an active research area. The recent growth in artificial intelligence software capable of processing complex language may provide new and more nuanced methods for similar tasks in future.

Clinical Validation

Validation then proceeded to clinical groups, but there were a number of recruitment challenges and too few participants were recruited to permit any analyses. The impacts of the Covid-19 pandemic on the health system may have played a role in the low referral rate. Hospital staff were under increased stress dealing with disruptions, changes to operating procedures, and significant staff shortages. There were fewer awake craniotomy surgeries conducted and the department's Neuropsychologist position was vacant during much of the study. Further, most referrals were made during the first half of the recruitment period, suggesting the study could have been forgotten amongst other demands on clinicians' time. It is also possible staff felt more protective of these patients due to their vulnerability, leading to hesitation about discussing study participation with them. Perhaps the detailed information sheets used for the study made participation appear more demanding than it actually was, especially for the community organisations who relied entirely on these for information about the study. Finally, the researcher found it difficult to balance the needs of the study with other demands of the clinical training programme and maintain momentum, delaying some communications and follow-up of referrals.

Future Research

The new test should not be used clinically without evidence for its validity. Establishing this will require further study by other researchers, perhaps in a wider international study. A similar study design may still be effective but with increased attention to reducing barriers to recruitment. If the researcher had spent more time on site at Wellington Regional Hospital it

may have been possible to pro-actively review admissions and/or approach potential participants directly, significantly reducing the demands on busy medical staff. The researcher's presence would also have served as a reminder the study was still recruiting, given staff the opportunity to discuss referrals they felt unsure about, and prevented delays in communication. Similar personal working relationships could have been developed with other hospitals and the community organisations approached for recruitment (e.g., attending regular meetings), rather than merely providing them with advertising material. The advertising material could also have involved summarised or otherwise simplified information sheets, with the detailed versions used only after potential participants expressed interest.

Conducting trials during awake craniotomy surgery, as originally planned, would be another way to examine the validity of the new test. Such trials would permit testing directly whether IFOF stimulation impairs test performance alongside establishing the test's suitability for use in a surgical context. A within-subjects design comparing performance with and without stimulation may also allow some conclusions to be drawn with a smaller sample size.

Conclusion

In summary, a new test was developed specifically for use during awake craniotomy, based on theoretical and clinical understandings of semantic association. Analyses of item quality and the test's performance with healthy adults suggested it improved upon existing tests and achieved its initial design objectives. While its validity could not be determined with this study's data, the test and its general population norms will be available for further research and then clinical use if indicated.

References

Adlam, A.-L. R., Patterson, K., Bozeat, S., & Hodges, J. R. (2010). The Cambridge Semantic Memory Test Battery: Detection of semantic deficits in semantic dementia and Alzheimer's disease. *Neurocase*, *16*(3), 193–207. https://doi.org/10.1080/13554790903405693

Agency for Clinical Innovation. (2019). *Telehealth in practice*. https://www.telehealth.org.nz/assets/standards/ACI-0261-Telehealth-guidelines.pdf

Alexander, H., Irwin, C., Purdie, G., & Hunn, M. (2010). Incidence and management of high grade glioma in Māori and non-Māori patients. *Journal of Clinical Neuroscience*, *17*(9), 1144–1147. https://doi.org/10.1016/j.jocn.2010.01.033

Almairac, F., Herbet, G., Moritz-Gasser, S., de Champfleur, N. M., & Duffau, H. (2015). The left inferior fronto-occipital fasciculus subserves language semantics: A multilevel lesion study. *Brain Structure and Function*, *220*(4), 1983–1995. https://doi.org/10.1007/s00429-014-0773-1

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, & Joint Committee on Standards for Educational

and Psychological Testing (U.S.) (Eds.). (2014). *Standards for educational and psychological testing*. American Educational Research Association.

American Psychological Association. (2020). *Informed consent checklist for telepsychological services*. https://www.apa.org/practice/programs/dmhi/research-information/informed-consent-checklist

Ardila, A., Ostrosky-Solís, F., & Bernal, B. (2006). Cognitive testing toward the future: The example of Semantic Verbal Fluency (ANIMALS). *International Journal of Psychology*, *41*(5), 324–332. https://doi.org/10.1080/00207590500345542

Bajada, C. J., Lambon Ralph, M. A., & Cloutman, L. L. (2015). Transport for language south of the Sylvian fissure: The routes and history of the main tracts and stations in the ventral language network. *Cortex*, *69*, 141–151. https://doi.org/10.1016/j.cortex.2015.05.011

Baradel, R. R., Silva, H. S. da, Estequi, J. G., Parente, M. A. de M. P., Sato, J. R., & Carthery-Goulart, M. T. (2014). Semantic memory for actions as assessed by the Kissing and Dancing Test: Education and age effects in cognitively healthy individuals. *Dementia & Neuropsychologia*, *8*(3), 216–222. https://doi.org/10.1590/S1980-57642014DN83000004

Berger, M. S., Kincaid, J., Ojemann, G. A., & Lettich, E. (1989). Brain Mapping Techniques to Maximize Resection, Safety, and Seizure Control in Children with Brain Tumors. *Neurosurgery*, *25*(5), 786–792. https://doi.org/10.1227/00006123-198911000-00015

Bernard, F., Clavreul, A., Casanova, M., Besnard, J., Lemée, J.-M., Soulard, G., Séguier, R., & Menei, P. (2023). Virtual Reality-Assisted Awake Craniotomy: A Retrospective Study. *Cancers*, *15*(3), 949. https://doi.org/10.3390/cancers15030949

Bilder, R. M., Postal, K. S., Barisa, M., Aase, D. M., Cullum, C. M., Gillaspy, S. R., Harder, L., Kanter, G., Lanca, M., Lechuga, D. M., Morgan, J. M., Most, R., Puente, A. E., Salinas, C. M., & Woodhouse, J. (2020). InterOrganizational practice committee recommendations/guidance for teleneuropsychology (TeleNP) in response to the COVID-19 pandemic. *The Clinical Neuropsychologist*, *34*(7–8), 1314–1334. https://doi.org/10.1080/13854046.2020.1767214

Bozdemir, M., & Gurvit, I. (2021). Normative data for the Turkish version of the pyramids and palm trees test. *Applied Neuropsychology: Adult*, *29*. https://doi.org/10.1080/23279095.2021.1900187

Brearly, T. W., Shura, R. D., Martindale, S. L., Lazowski, R. A., Luxton, D. D., Shenal, B. V., & Rowland, J. A. (2017). Neuropsychological Test Administration by Videoconference: A Systematic Review and Meta-Analysis. *Neuropsychology Review*, *27*(2), 174–186. https://doi.org/10.1007/s11065-017-9349-1

Breining, B. L., Lala, T., Cuitiño, M. M., Manes, F., Peristeri, E., Tsapkini, K., Faria, A. V., & Hillis, A. E. (2015). A brief assessment of object semantics in primary progressive aphasia. *Aphasiology*, *29*(4), 488–505. https://doi.org/10.1080/02687038.2014.973360

Brown, T. J., Brennan, M. C., Li, M., Church, E. W., Brandmeir, N. J., Rakszawski, K. L., Patel, A. S., Rizk, E. B., Suki, D., Sawaya, R., & Glantz, M. (2016). Association of the Extent of Resection With Survival in Glioblastoma: A Systematic Review and Meta-analysis. *JAMA Oncology*, *2*(11), 1460–1469. https://doi.org/10.1001/jamaoncol.2016.1373

Callahan, B. L., Macoir, J., Hudon, C., Bier, N., Chouinard, N., Cossette-Harvey, M., Daigle, N., Fradette, C., Gagnon, L., & Potvin, O. (2010). Normative Data for the Pyramids and Palm

Trees Test in the Quebec-French Population. *Archives of Clinical Neuropsychology*, *25*(3), 212–217. https://doi.org/10.1093/arclin/acq013

Cejudo, J. C., Samaniego, M., Almeria, M., Castrillo, S., Medina, L., & Gil, D. (2022). Ikos Test: New Tool for the Assessment of Semantic Knowledge in Early Alzheimer Disease. *Journal of Alzheimer's Disease*, *90*(1), 151–160. https://doi.org/10.3233/JAD-220516

Chang, W.-H., Pei, Y.-C., Wei, K.-C., Chao, Y.-P., Chen, M.-H., Yeh, H.-A., Jaw, F.-S., & Chen, P.-Y. (2018). Intraoperative linguistic performance during awake brain surgery predicts postoperative linguistic deficits. *Journal of Neuro-Oncology*, *139*(1), 215–223. https://doi.org/10.1007/s11060-018-2863-z

Chen, H.-F., Huang, L.-L., Li, H.-Y., Qian, Y., Yang, D., Qing, Z., Luo, C.-M., Li, M.-C., Zhang, B., & Xu, Y. (2020). Microstructural disruption of the right inferior fronto-occipital and inferior longitudinal fasciculus contributes to WMH-related cognitive impairment. *CNS Neuroscience & Therapeutics*, *26*(5), 576–588. https://doi.org/10.1111/cns.13283

Chen, L., Lambon Ralph, M. A., & Rogers, T. T. (2017). A unified model of human semantic knowledge and its disorders. *Nature Human Behaviour*, *1*(3), 0039. https://doi.org/10.1038/s41562-016-0039

Chen, L., & Rogers, T. T. (2014). Revisiting domain-general accounts of category specificity in mind and brain. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(3), 327–344. https://doi.org/10.1002/wcs.1283

Chua, T. H., See, A. A. Q., Ang, B. T., & King, N. K. K. (2018). Awake Craniotomy for Resection of Brain Metastases: A Systematic Review. *World Neurosurgery*, *120*, e1128–e1135. https://doi.org/10.1016/j.wneu.2018.08.243

Coello, A. F., Moritz-Gasser, S., Martino, J., Martinoni, M., Matsuda, R., & Duffau, H. (2013). Selection of intraoperative tasks for awake mapping based on relationships between tumor location and functional networks: A review. *Journal of Neurosurgery*, *119*(6), 1380–1394. https://doi.org/10.3171/2013.6.JNS122470

Conner, A. K., Briggs, R. G., Sali, G., Rahimi, M., Baker, C. M., Burks, J. D., Glenn, C. A., Battiste, J. D., & Sughrue, M. E. (2018). A Connectomic Atlas of the Human Cerebrum—Chapter 13: Tractographic Description of the Inferior Fronto-Occipital Fasciculus. *Operative Neurosurgery*, *15*(suppl_1), S436–S443. https://doi.org/10.1093/ons/opy267

Cormack, D., Robson, C., & Te Rōpū Rangahau Hauora a Eru Pōmare. (2011). *Classification and output of multiple ethnicities: Considerations for monitoring Māori health*. Te Rōpū Rangahau Hauora a Eru Pōmare. http://www.ethnicity.maori.nz/files/Mulitple_Ethnicities_WEB_ISBN.pdf

Coyle-Gilchrist, I. T. S., Dick, K. M., Patterson, K., Vázquez Rodríquez, P., Wehmann, E., Wilcox, A., Lansdall, C. J., Dawson, K. E., Wiggins, J., Mead, S., Brayne, C., & Rowe, J. B. (2016). Prevalence, characteristics, and survival of frontotemporal lobar degeneration syndromes. *Neurology*, *86*(18), 1736–1743. https://doi.org/10.1212/WNL.0000000000002638

Davis, M. E. (2018). Epidemiology and Overview of Gliomas. *Seminars in Oncology Nursing*, *34*(5), 420–429. https://doi.org/10.1016/j.soncn.2018.10.001

De Witt Hamer, P., Gil-Robles, S., Zwinderman, A., Duffau, H., & Berger, M. (2012). Impact of Intraoperative Stimulation Brain Mapping on Glioma Surgery Outcome: A Meta-Analysis.

Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology, *30*, 2559–2565. https://doi.org/10.1200/JCO.2011.38.4818

De Witte, E., Satoer, D., Robert, E., Colle, H., Verheyen, S., Visch-Brink, E., & Mariën, P. (2015). The Dutch Linguistic Intraoperative Protocol: A valid linguistic approach to awake brain surgery. *Brain and Language*, *140*, 35–48. https://doi.org/10.1016/j.bandl.2014.10.011

Duffau, H., Herbet, G., & Moritz-Gasser, S. (2013). Toward a pluri-component, multimodal, and dynamic organization of the ventral semantic stream in humans: Lessons from stimulation mapping in awake patients. *Frontiers in Systems Neuroscience*, *7*. https://doi.org/10.3389/fnsys.2013.00044

Duffau, H., Moritz-Gasser, S., & Mandonnet, E. (2014). A re-examination of neural basis of language processing: Proposal of a dynamic hodotopical model from data provided by brain stimulation mapping during picture naming. *Brain and Language*, *131*, 1–10. https://doi.org/10.1016/j.bandl.2013.05.011

Fares, M., Kutuzov, A., Oepen, S., & Velldal, E. (2017). Word vectors, reuse, and replicability: Towards a community repository of large-text resources. In J. Tiedemann (Ed.), *Proceedings of the 21st Nordic Conference of Computational Linguistics, NoDaLiDa, 22-24 May 2017.* (pp. 271–276). Linkoping University Electronic Press.

Fergadiotis, G., Wright, H., & Capilouto, G. (2010). Psychometric Properties of the Pyramids and Palm Trees Test. *Procedia - Social and Behavioral Sciences*, *6*, 33–34. https://doi.org/10.1016/j.sbspro.2010.08.017

Gamboz, N., Coluccia, E., Iavarone, A., & Brandimonte, M. A. (2009). Normative data for the Pyramids and Palm Trees Test in the elderly Italian population. *Neurological Sciences*, *30*(6), 453–458. https://doi.org/10.1007/s10072-009-0130-y

Garrard, P., & Carroll, E. (2006). Lost in semantic space: A multi-modal, non-verbal assessment of feature knowledge in semantic dementia. *Brain*, *129*(5), 1152–1163. https://doi.org/10.1093/brain/awl069

Gernsback, J. E., Kolcun, J. P. G., Starke, R. M., Ivan, M. E., & Komotar, R. J. (2018). Who Needs Sleep? An Analysis of Patient Tolerance in Awake Craniotomy. *World Neurosurgery*, *118*, e842–e848. https://doi.org/10.1016/j.wneu.2018.07.064

Gerritsen, J. K. W., Zwarthoed, R. H., Kilgallon, J. L., Nawabi, N. L., Jessurun, C. A. C., Versyck, G., Pruijn, K. P., Fisher, F. L., Larivière, E., Solie, L., Mekary, R. A., Satoer, D. D., Schouten, J. W., Bos, E. M., Kloet, A., Nandoe Tewarie, R., Smith, T. R., Dirven, C. M. F., De Vleeschouwer, S., … Vincent, A. J. P. E. (2022). Effect of awake craniotomy in glioblastoma in eloquent areas (GLIOMAP): A propensity score-matched analysis of an international, multicentre, cohort study. *The Lancet Oncology*, *23*(6), 802–817. https://doi.org/10.1016/S1470-2045(22)00213-3

Gogos, A. J., Young, J. S., Morshed, R. A., Hervey-Jumper, S. L., & Berger, M. S. (2020). Awake glioma surgery: Technical evolution and nuances. *Journal of Neuro-Oncology*, *147*(3), 515–524. https://doi.org/10.1007/s11060-020-03482-z

Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., Ogar, J. M., Rohrer, J. D., Black, S., Boeve, B. F., Manes, F., Dronkers, N. F., Vandenberghe, R., Rascovsky, K., Patterson, K., Miller, B. L., Knopman, D. S., Hodges, J. R., Mesulam, M. M.,

& Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, *76*(11), 1006–1014. https://doi.org/10.1212/WNL.0b013e31821103e6

Government Statistical Service. (2015). *Primary Principles: Ethnic Group* (Harmonised Concepts and Questions for Social Data Sources). https://gss.civilservice.gov.uk/wp-content/uploads/2016/03/P3-Ethnic-Group-June-16-1.pdf

Gudayol-Ferré, E., Lara, J. P., Herrera-Guzman, I., Böhm, P., Rodés, E., Ansaldo, A. I., & Peña-Casanova, J. (2008). Semantic memory as assessed by the Pyramids and Palm Trees Test: The impact of sociodemographic factors in a Spanish-speaking population. *Journal of the International Neuropsychological Society*, *14*(01). https://doi.org/10.1017/S1355617708080168

Guo, Q., He, C., Wen, X., Song, L., Han, Z., & Bi, Y. (2014). Adapting the Pyramids and Palm Trees Test and the Kissing and Dancing Test and developing other semantic tests for the Chinese population. *Applied Psycholinguistics; New York*, *35*(6), 1001–1019. http://dx.doi.org.ezproxy.massey.ac.nz/10.1017/S0142716412000677

Haladyna, T. M. (2015). Item Analysis for Selected-Response Test Items. In S. Lane, M. R. Raymond, & T. M. Haladyna, *Handbook of Test Development* (Second edition). Routledge. https://doi.org/10.4324/9780203102961-28

Han, Z., Ma, Y., Gong, G., He, Y., Caramazza, A., & Bi, Y. (2013). White matter structural connectivity underlying semantic processing: Evidence from brain damaged patients. *Brain*, *136*(10), 2952–2965. https://doi.org/10.1093/brain/awt205

Hande, V., Gunasekaran, H., Hegde, S., Shashidhar, A., & Arimappamagan, A. (2021). Role of Clinical Neuropsychologists in Awake-Craniotomy. *Neurology India*, *69*(3), 711. https://doi.org/10.4103/0028-3886.319237

Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, *7*(3), 238–247. https://doi.org/10.1037/1040-3590.7.3.238

Henry, J. D., & Crawford, J. R. (2004). A Meta-Analytic Review of Verbal Fluency Performance Following Focal Cortical Lesions. *Neuropsychology*, *18*(2), 284–295. https://doi.org/10.1037/0894-4105.18.2.284

Herbet, G., Moritz-Gasser, S., & Duffau, H. (2017). Direct evidence for the contributive role of the right inferior fronto-occipital fasciculus in non-verbal semantic cognition. *Brain Structure and Function*, *222*(4), 1597–1610. https://doi.org/10.1007/s00429-016-1294-x

Heuer, S. (2016). The influence of image characteristics on image recognition: A comparison of photographs and line drawings. *Aphasiology*, *30*(8), 943–961. https://doi.org/10.1080/02687038.2015.1081138

Heuer, S., & Hallowell, B. (2009). Visual attention in a multiple-choice task: Influences of image characteristics with and without presentation of a verbal stimulus. *Aphasiology*, *23*(3), 351–363. https://doi.org/10.1080/02687030701770474

Hogan, D. B., Jetté, N., Fiest, K. M., Roberts, J. I., Pearson, D., Smith, E. E., Roach, P., Kirk, A., Pringsheim, T., & Maxwell, C. J. (2016). The Prevalence and Incidence of Frontotemporal Dementia: A Systematic Review. *Canadian Journal of Neurological Sciences*, *43*(S1), S96–S109. https://doi.org/10.1017/cjn.2016.25

Howard, D., & Patterson, K. (1992). *The Pyramids and Palm Trees Test* [Measurement Instrument]. Pearson Assessment.

Irish, M., & Vatansever, D. (2020). Rethinking the episodic-semantic distinction from a gradient perspective. *Current Opinion in Behavioral Sciences*, *32*, 43–49. https://doi.org/10.1016/j.cobeha.2020.01.016

Joint Task Force for the Development of Telepsychology Guidelines for Psychologists. (2013). Guidelines for the practice of telepsychology. *American Psychologist*, *68*(9), 791–800. https://doi.org/10.1037/a0035001

Kaplan, E., Goodglass, H., Weintraub, S., & Goodglass, H. (1983). *Boston naming test*. Lea & Febiger.

Klein, L. A., & Buchanan, J. A. (2009). Psychometric properties of the Pyramids and Palm Trees Test. *Journal of Clinical and Experimental Neuropsychology*, *31*(7), 803–808. https://doi.org/10.1080/13803390802508926

Landauer, T. K., & Dutnais, S. T. (1997). A Solution to Plato's Problem: The Latent Semantic Analysis Theory of Acquisition, Induction, and Representation of Knowledge. *Psychological Review*, *104*(2), 30.

Leeds and York Partnership. (2021). *Easy on the i*. Learning Disability Service. https://www.learningdisabilityservice-leeds.nhs.uk/easy-on-the-i/image-bank/

Lemée, J.-M., Bernard, F., Ter Minassian, A., & Menei, P. (2018). Right Hemisphere Cognitive Functions: From Clinical and Anatomical Bases to Brain Mapping During Awake Craniotomy. Part II: Neuropsychological Tasks and Brain Mapping. *World Neurosurgery*, *118*, 360–367. https://doi.org/10.1016/j.wneu.2018.07.099

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764–766. https://doi.org/10.1016/j.jesp.2013.03.013

Louis, D. N., Ohgaki, H., Wiestler, O. D., Cavenee, W. K., Burger, P. C., Jouvet, A., Scheithauer, B. W., & Kleihues, P. (2007). The 2007 WHO Classification of Tumours of the Central Nervous System. *Acta Neuropathologica*, *114*(2), 97–109. https://doi.org/10.1007/s00401-007-0243-4

Luzzatti, C., Mauri, I., Castiglioni, S., Zuffi, M., Spartà, C., Somalvico, F., & Franceschi, M. (2020). Evaluating Semantic Knowledge Through a Semantic Association Task in Individuals With Dementia. *American Journal of Alzheimer's Disease & Other Dementias®*, *35*, 1533317520917294. https://doi.org/10.1177/1533317520917294

Mandonnet, E., Sarubbo, S., & Duffau, H. (2017). Proposal of an optimized strategy for intraoperative testing of speech and language during awake mapping. *Neurosurgical Review*, *40*(1), 29–35. https://doi.org/10.1007/s10143-016-0723-x

Marshall, C. R., Hardy, C. J. D., Volkmer, A., Russell, L. L., Bond, R. L., Fletcher, P. D., Clark, C. N., Mummery, C. J., Schott, J. M., Rossor, M. N., Fox, N. C., Crutch, S. J., Rohrer, J. D., & Warren, J. D. (2018). Primary progressive aphasia: A clinical approach. *Journal of Neurology*, *265*(6), 1474–1490. https://doi.org/10.1007/s00415-018-8762-6

Martin, A. (2016). GRAPES—Grounding representations in action, perception, and emotion systems: How object properties and categories are represented in the human brain.

Psychonomic Bulletin & Review, *23*(4), 979–990. https://doi.org/10.3758/s13423-015-0842-3

Martínez-Cuitiño, M., & Barreyro, J. P. (2010). ¿PIRÁMIDES Y PALMERAS O PIRÁMIDES Y FARAONES? ADAPTACIÓN Y VALIDACIÓN DE UN TEST DE ASOCIACIÓN SEMÁNTICA AL ESPAÑOL RIOPLATENSE. *Interdisciplinaria*, *27*(2). http://www.redalyc.org/resumen.oa?id=18018446004

Matthews, K., Phelan, J., Jones, N. A., Konya, S., Marks, R., Pratt, B., Coombs, J., & Bentley, M. (2017). *2015 National Content Test: Race and Ethnicity Analysis Report* (p. 380). United States Census Bureau.

Mazoué, A., Gaultier, A., Rocher, L., Deruet, A.-L., Vercelletto, M., & Boutoleau-Bretonnière, C. (2022). Does a rabbit have feathers or fur? Development of a 42-item semantic memory test (SMT-42). *Journal of Clinical and Experimental Neuropsychology*, *44*(7), 514–531. https://doi.org/10.1080/13803395.2022.2133088

Mazoyer, V., Knoblauch, K., Fontanay, S., Koenig, F., & Vital-Durand, F. (1999). Identification of low-pass filtered images by low vision patients. *Visual Impairment Research*, *1*(3), 187–192. https://doi.org/10.1076/vimr.1.3.187.4441

McHugh, M. L. (2013). The Chi-square test of independence. *Biochemia Medica*, 143–149. https://doi.org/10.11613/BM.2013.018

Mehri, A., Mousavi, S. Z., Kamali, M., & Maroufizadeh, S. (2018). Normative data for the Pyramids and Palm Trees Test in literate Persian adults. *Iranian Journal of Neurology*, *17*(1), 18–23.

Mofatteh, M., Mashayekhi, M. S., Arfaie, S., Chen, Y., Hendi, K., Kwan, A. T. H., Honarvar, F., Solgi, A., Liao, X., & Ashkan, K. (2023). Stress, Anxiety, and Depression Associated With Awake Craniotomy: A Systematic Review. *Neurosurgery*, *92*(2), 225–240. https://doi.org/10.1227/neu.0000000000002224

Moritz-Gasser, S., Herbet, G., & Duffau, H. (2013). Mapping the connectivity underlying multimodal (verbal and non-verbal) semantic processing: A brain electrostimulation study. *Neuropsychologia*, *51*(10), 1814–1822. https://doi.org/10.1016/j.neuropsychologia.2013.06.007

National Ethics Advisory Committee. (2019). *National Ethical Standards for Health and Disability Research and Quality Improvement*. Ministry of Health.

Neill, E., Gurvich, C., & Rossell, S. L. (2014). Category fluency in schizophrenia research: Is it an executive or semantic measure? *Cognitive Neuropsychiatry*, *19*(1), 81–95. https://doi.org/10.1080/13546805.2013.807233

Ohlerth, A.-K., Valentin, A., Vergani, F., Ashkan, K., & Bastiaanse, R. (2019). The verb and noun test for peri-operative testing (VAN-POP): Standardized language tests for navigated transcranial magnetic stimulation and direct electrical stimulation. *Acta Neurochirurgica*. https://doi.org/10.1007/s00701-019-04159-x

Ohman, A., Sheppard, C., Monetta, L., & Taler, V. (2020). Assessment of semantic memory in mild cognitive impairment: The psychometric properties of a novel semantic battery. *Applied Neuropsychology: Adult*, *0*(0), 1–7. https://doi.org/10.1080/23279095.2020.1774885

Olivarez, A., & Boroda, A. D. (2007). Review of The Pyramids and Palm Trees Test. In K. F. Geisinger, R. A. Spies, J. F. Carlson, & B. S. Plake (Eds.), *The seventeenth mental measurements yearbook*.

Pallud, J., Mandonnet, E., Corns, R., Dezamis, E., Parraga, E., Zanello, M., & Spena, G. (2017). Technical principles of direct bipolar electrostimulation for cortical and subcortical mapping in awake craniotomy. *Neurochirurgie*, *63*(3), 158–163. https://doi.org/10.1016/j.neuchi.2016.12.004

Paplikar, A., Vandana, V. P., Mekala, S., Darshini, K. J., Arshad, F., Iyer, G. K., Kandukuri, R., Divyaraj, G., Varghese, F., Kaul, S., Patterson, K., & Alladi, S. (2022). Semantic memory impairment in dementia: A cross-cultural adaptation study. *Neurological Sciences*, *43*(1), 265–273. https://doi.org/10.1007/s10072-021-05272-5

Patterson, K., Kopelman, M. D., Woollams, A. M., Brownsett, S. L. E., Geranmayeh, F., & Wise, R. J. S. (2015). Semantic memory: Which side are you on? *Neuropsychologia*, *76*, 182–191. https://doi.org/10.1016/j.neuropsychologia.2014.11.024

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, *8*(12), 976–987. https://doi.org/10.1038/nrn2277

Penfield, W., & Boldrey, E. (1937). SOMATIC MOTOR AND SENSORY REPRESENTATION IN THE CEREBRAL CORTEX OF MAN AS STUDIED BY ELECTRICAL STIMULATION. *Brain*, *60*(4), 389–443. https://doi.org/10.1093/brain/60.4.389

Penfield, W., & Roberts, L. (1959). *Speech and Brain Mechanisms*. Princeton University Press.

People First New Zealand. (2014). *A guide to making Easy Read information*. https://www.odi.govt.nz/assets/Guidance-and-Resources-files/guide-to-writing-easy-read-information-people-first-2014.pdf

Pozueta, A., Lage, C., Martínez, M. G., Kazimierczak, M., Bravo, M., López-García, S., Riancho, J., González-Suarez, A., Vázquez-Higuera, J. L., de Arcocha-Torres, M., Banzo, I., Bonilla, J. J., Berciano, J., Rodríguez-Rodríguez, E., & Sánchez-Juan, P. (2019). A Brief Drawing Task for the Differential Diagnosis of Semantic Dementia. *Journal of Alzheimer's Disease*, *72*(1), 151–160. https://doi.org/10.3233/JAD-190660

Price, A. R., Bonner, M. F., & Grossman, M. (2015). Semantic Memory: Cognitive and Neuroanatomical Perspectives. In *Brain Mapping* (pp. 529–536). Elsevier. https://doi.org/10.1016/B978-0-12-397025-1.00280-3

Ralph, M. A. L., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, *18*(1), 42–55. https://doi.org/10.1038/nrn.2016.150

Rami, L., Serradell, M., Bosch, B., Caprile, C., Sekler, A., Villar, A., Canal, R., & Molinuevo, J. L. (2008). Normative data for the Boston Naming Test and the Pyramids and Palm Trees Test in the elderly Spanish population. *Neuropsychology, Development, and Cognition. Section A, Journal of Clinical and Experimental Neuropsychology*, *1*, 1.

RANDOM.ORG. (2021). *RANDOM.ORG - True Random Number Service*. https://www.random.org/

Řehůřek, R., & Sojka, P. (2010). Software Framework for Topic Modelling with Large Corpora. *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, 45–50.

Rey, S. (2020). *Développement et faisabilité d'un test de dessin de mémoire chez des patients suspectés de présenter un trouble sémantique* [Master's Thesis, Université de Rouen]. Dépôt Universitaire de Mémoires Après Soutenance. https://dumas.ccsd.cnrs.fr/dumas-02903449

Ries, S. K., Piai, V., Perry, D., Griffin, S., Jordan, K., Henry, R., Knight, R. T., & Berger, M. S. (2019). Roles of ventral versus dorsal pathways in language production: An awake language mapping study. *Brain and Language*, *191*, 17–27. https://doi.org/10.1016/j.bandl.2019.01.001

Rofes, A., Spena, G., Miozzo, A., Fontanella, M. M., & Miceli, G. (2015). Advantages and disadvantages of intraoperative language tasks in awake surgery: A three-task approach for prefrontal tumors. *JOURNAL OF NEUROSURGICAL SCIENCES*, *59*(4), 15.

Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., & Patterson, K. (2004). Structure and Deterioration of Semantic Memory: A Neuropsychological and Computational Investigation. *Psychological Review*, *111*(1), 205–235. https://doi.org/10.1037/0033-295X.111.1.205

Rogers, T. T., Lambon Ralph, M. A., Hodges, J. R., & Patterson, K. (2004). Natural Selection: The Impact of Semantic Impairment on Lexical and Object Decision. *Cognitive Neuropsychology*, *21*(2–4), 331–352. https://doi.org/10.1080/02643290342000366

Rollans, C., & Cummine, J. (2018). One tract, two tract, old tract, new tract: A pilot study of the structural and functional differentiation of the inferior fronto-occipital fasciculus. *Journal of Neurolinguistics*, *46*, 122–137. https://doi.org/10.1016/j.jneuroling.2017.12.009

Rosenmai, P. (2013). *Using the Median Absolute Deviation to Find Outliers*. Eureka Statistics. https://eurekastatistics.com/using-the-median-absolute-deviation-to-find-outliers/

Ruis, C. (2018). Monitoring cognition during awake brain surgery in adults: A systematic review. *Journal of Clinical and Experimental Neuropsychology*, *40*(10), 1081–1104. https://doi.org/10.1080/13803395.2018.1469602

Schmidt, C. S. M., Schumacher, L. V., Römer, P., Leonhart, R., Beume, L., Martin, M., Dressing, A., Weiller, C., & Kaller, C. P. (2017). Are semantic and phonological fluency based on the same or distinct sets of cognitive processes? Insights from factor analyses in healthy adults and stroke patients. *Neuropsychologia*, *99*, 148–155. https://doi.org/10.1016/j.neuropsychologia.2017.02.019

Scott, G. G., Keitel, A., Becirspahic, M., Yao, B., & Sereno, S. C. (2018). The Glasgow Norms: Ratings of 5,500 words on nine scales. *Behavior Research Methods*. https://doi.org/10.3758/s13428-018-1099-3

Sharpe, D. (2015). *Chi-Square Test is Statistically Significant: Now What? 20*(8), 11.

Šimkovic, M., & Träuble, B. (2019). Robustness of statistical methods when measure is affected by ceiling and/or floor effect. *PLOS ONE*, *14*(8), e0220889. https://doi.org/10.1371/journal.pone.0220889

Singh, P. (2014). 2 x 2 Contingency Table: Fisher's Exact Test. *Journal of Clinical and Preventive Cardiology*, *2*.

Statistics New Zealand. (2021). *Statistical standard for ethnicity v1.0.0*. http://www.stats.govt.nz

Stockbridge, M. D., Tippett, D. C., Breining, B. L., Vitti, E., & Hillis, A. E. (2021). Task performance to discriminate among variants of primary progressive aphasia. *Cortex*, *145*, 201–211. https://doi.org/10.1016/j.cortex.2021.09.015

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (Manawatu 616.80475 Str; 3rd ed). Oxford University Press.

Suarez-Meade, P., Marenco-Hillembrand, L., Prevatt, C., Murguia-Fuentes, R., Mohamed, A., Alsaeed, T., Lehrer, E. J., Brigham, T., Ruiz-Garcia, H., Sabsevitz, D., Middlebrooks, E. H., Bechtle, P. S., Quinones-Hinojosa, A., & Chaichana, K. L. (2020). Awake vs. asleep motor mapping for glioma resection: A systematic review and meta-analysis. *Acta Neurochirurgica*, *162*(7), 1709–1720. https://doi.org/10.1007/s00701-020-04357-y

Surbeck, W., Hänggi, J., Scholtes, F., Viher, P. V., Schmidt, A., Stegmayer, K., Studerus, E., Lang, U. E., Riecher-Rössler, A., Strik, W., Seifritz, E., Borgwardt, S., Quednow, B. B., & Walther, S. (2020). Anatomical integrity within the inferior fronto-occipital fasciculus and semantic processing deficits in schizophrenia spectrum disorders. *Schizophrenia Research*, *218*, 267–275. https://doi.org/10.1016/j.schres.2019.12.025

Talacchi, A., Santini, B., Casagrande, F., Alessandrini, F., Zoccatelli, G., & Squintani, G. M. (2013). Awake surgery between art and science. Part I: Clinical and operative settings. *Functional Neurology*, *28*(3), 205–221.

Talacchi, A., Santini, B., Casartelli, M., Monti, A., Capasso, R., & Miceli, G. (2013). Awake surgery between art and science. Part II: Language and cognitive mapping. *Functional Neurology*, *28*(3), 223–239.

Tulving, E. (1972). 12. Episodic and Semantic Memory. In E. Tulving & W. Donaldson (Eds.), *Organization of Memory* (pp. 381–403). Academic Press.

UNESCO Institute for Statistics. (2012). *International standard classification of education: ISCED 2011*. UNESCO Institute for Statistics. http://www.uis.unesco.org/Education/Documents/isced-2011-en.pdf

van Heesewijk, J., Steenwijk, M. D., Kreukels, B. P. C., Veltman, D. J., Bakker, J., & Burke, S. M. (2022). Alterations in the inferior fronto-occipital fasciculus – a specific neural correlate of gender incongruence? *Psychological Medicine*, 1–10. https://doi.org/10.1017/S0033291721005547

Verst, S. M., de Castro, I., Scappini-Junior, W., de Melo, M. N., de Oliveira, J. R., de Almeida, S. S., Alvarez, N. R. C., Sucena, A. C. B., Barros, M. R., Marrone, C. D., & Maldaun, M. V. C. (2021). Methodology for creating and validating object naming and semantic tests used by Verst-Maldaun Language Assessment during awake craniotomies. *Clinical Neurology and Neurosurgery*, *202*, 106485. https://doi.org/10.1016/j.clineuro.2021.106485

Vilasboas, T., Herbet, G., & Duffau, H. (2017). Challenging the Myth of Right Nondominant Hemisphere: Lessons from Corticosubcortical Stimulation Mapping in Awake Surgery and Surgical Implications. *World Neurosurgery*, *103*, 449–456. https://doi.org/10.1016/j.wneu.2017.04.021

Wadsworth, H. E., Dhima, K., Womack, K. B., Hart, J., Weiner, M. F., Hynan, L. S., & Cullum, C. M. (2018). Validity of Teleneuropsychological Assessment in Older Patients with Cognitive Disorders. *Archives of Clinical Neuropsychology: The Official Journal of the*

National Academy of Neuropsychologists, *33*(8), 1040–1045. https://doi.org/10.1093/arclin/acx140

Welch, B. M., Marshall, E., Qanungo, S., Aziz, A., Laken, M., Lenert, L., & Obeid, J. (2016). Teleconsent: A novel approach to obtain informed consent for research. *Contemporary Clinical Trials Communications*, *3*, 74–79. https://doi.org/10.1016/j.conctc.2016.03.002

Weller, M., van den Bent, M., Preusser, M., Le Rhun, E., Tonn, J. C., Minniti, G., Bendszus, M., Balana, C., Chinot, O., Dirven, L., French, P., Hegi, M. E., Jakola, A. S., Platten, M., Roth, P., Rudà, R., Short, S., Smits, M., Taphoorn, M. J. B., … Wick, W. (2021). EANO guidelines on the diagnosis and treatment of diffuse gliomas of adulthood. *Nature Reviews Clinical Oncology*, *18*(3), Article 3. https://doi.org/10.1038/s41571-020-00447-z

Weller, M., Wick, W., Aldape, K., Brada, M., Berger, M., Pfister, S. M., Nishikawa, R., Rosenthal, M., Wen, P. Y., Stupp, R., & Reifenberger, G. (2015). Glioma. *Nature Reviews Disease Primers*, *1*(1), Article 1. https://doi.org/10.1038/nrdp.2015.17

Whitaker, H. A., & Ojemann, G. A. (1977). Graded localisation of naming from electrical stimulation mapping of left cerebral cortex. *Nature*, *270*(5632), Article 5632. https://doi.org/10.1038/270050a0

Xu, Y., He, Y., & Bi, Y. (2017). A Tri-network Model of Human Semantic Processing. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.01538

Yeh, F.-C., Panesar, S., Fernandes, D., Meola, A., Yoshino, M., Fernandez-Miranda, J. C., Vettel, J. M., & Verstynen, T. (2018). Population-averaged atlas of the macroscale human structural connectome and its network topology. *NeuroImage*, *178*, 57–68. https://doi.org/10.1016/j.neuroimage.2018.05.027

Zannino, G. D., Perri, R., Marra, C., Caruso, G., Baroncini, M., Caltagirone, C., & Carlesimo, G. A. (2021). The Free Association Task: Proposal of a Clinical Tool for Detecting Differential Profiles of Semantic Impairment in Semantic Dementia and Alzheimer's Disease. *Medicina*, *57*(11), Article 11. https://doi.org/10.3390/medicina57111171

Zhang, J. J. Y., Lee, K. S., Voisin, M. R., Hervey-Jumper, S. L., Berger, M. S., & Zadeh, G. (2020). Awake craniotomy for resection of supratentorial glioblastoma: A systematic review and meta-analysis. *Neuro-Oncology Advances*, *2*(vdaa111). https://doi.org/10.1093/noajnl/vdaa111

Zieky, M. J. (2015). Developing Fair Tests. In S. Lane, M. R. Raymond, & T. M. Haladyna, *Handbook of Test Development* (Second edition). Routledge. https://doi.org/10.4324/9780203102961-11

Appendix A: Final New Test

Administration instructions:

Show teaching item.

Say: **Here are three pictures. You have to decide which picture is correct. Which one of these do you think is *correct*?**

Say: [**Yes, that's right.** / **No, it's this one.**] **The sun and a cloud can appear together in the sky but neither pencils nor fridges are likely to be in a cloud.**

All the test questions are similar. For each question, choose which picture you think is *correct*. If you're not sure, take a guess.

Any questions?

Provide as much explanation as needed for the examinee to understand the task before proceeding to the test items.

Repeat the instruction "**Which picture is *correct*?**" as often as is necessary.

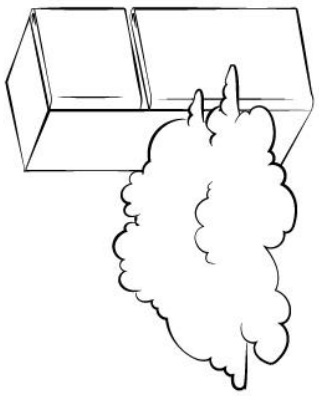
If examinee says multiple pictures could be correct, say: **Maybe, but which one is the *most likely*?**

If examinee asks what an object is, say: **Just answer based on what it looks like to you.**

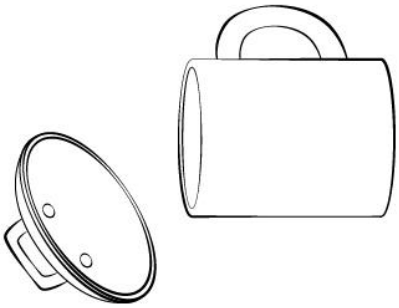
Scoring:

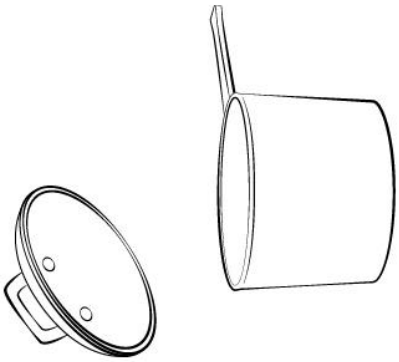
Award one point for each correct answer, giving a score from 0-58. Scores of 56 or less were very unusual in the norming sample so may indicate impairment.

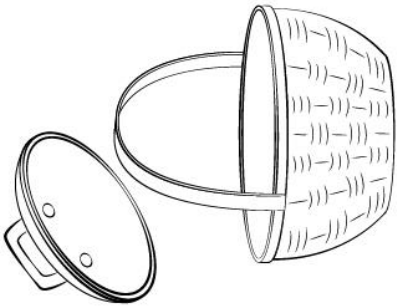
Teaching item



Item 1

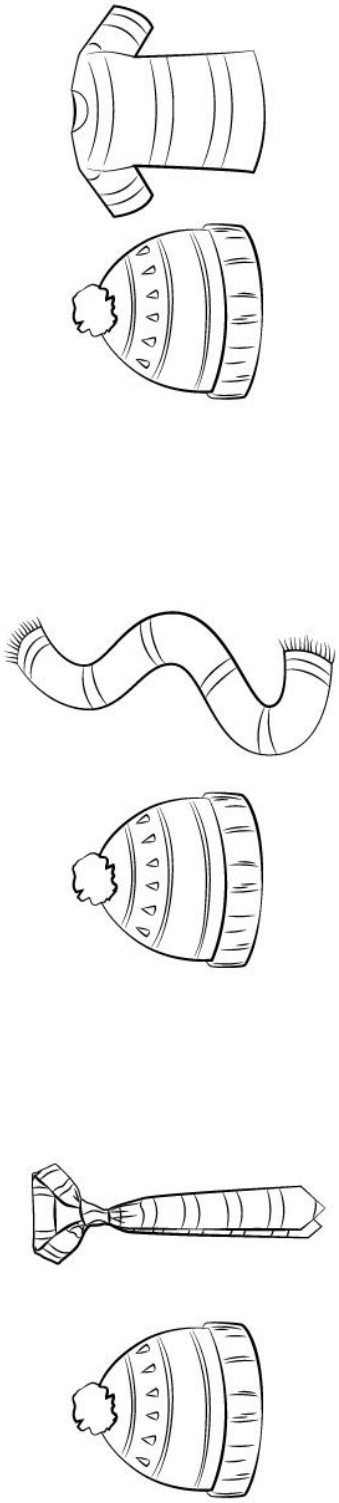




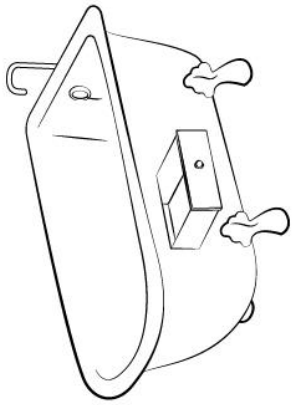


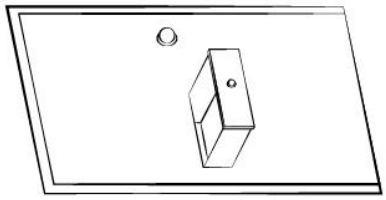
Item 2

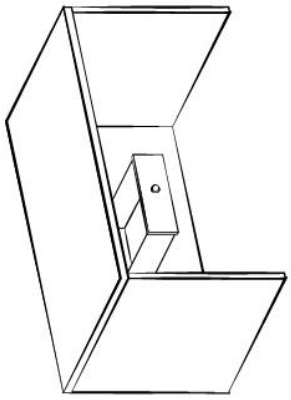
Item 3



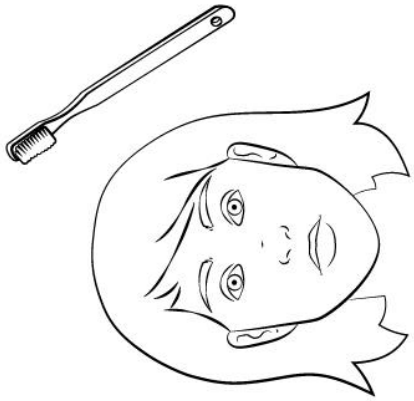
Item 4





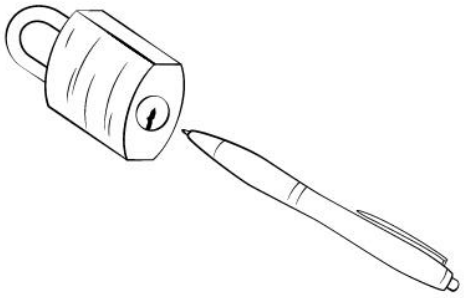


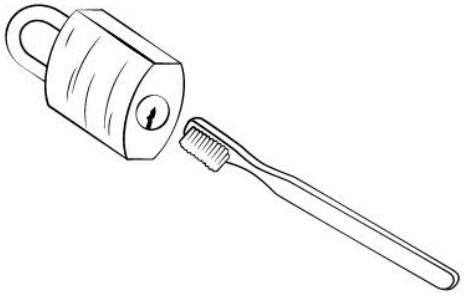
Item 5

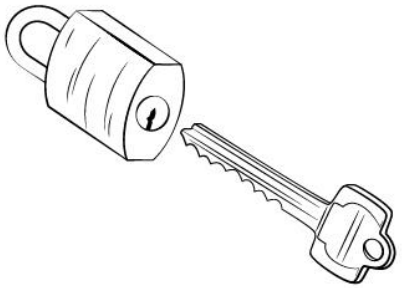


Item 6

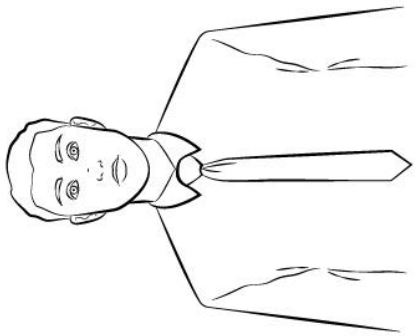
Item 7

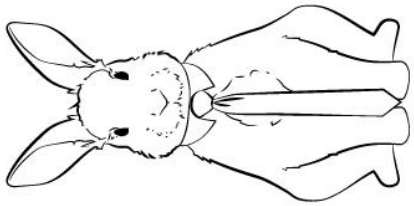


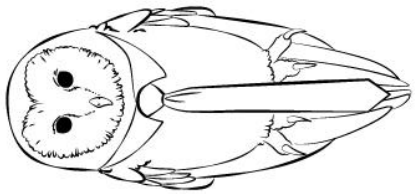




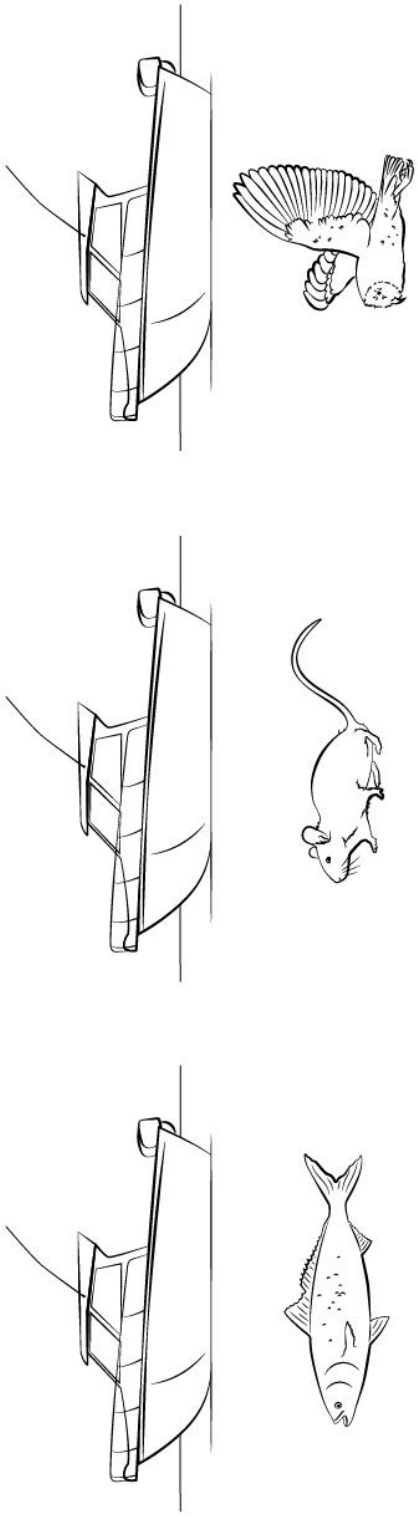
Item 8



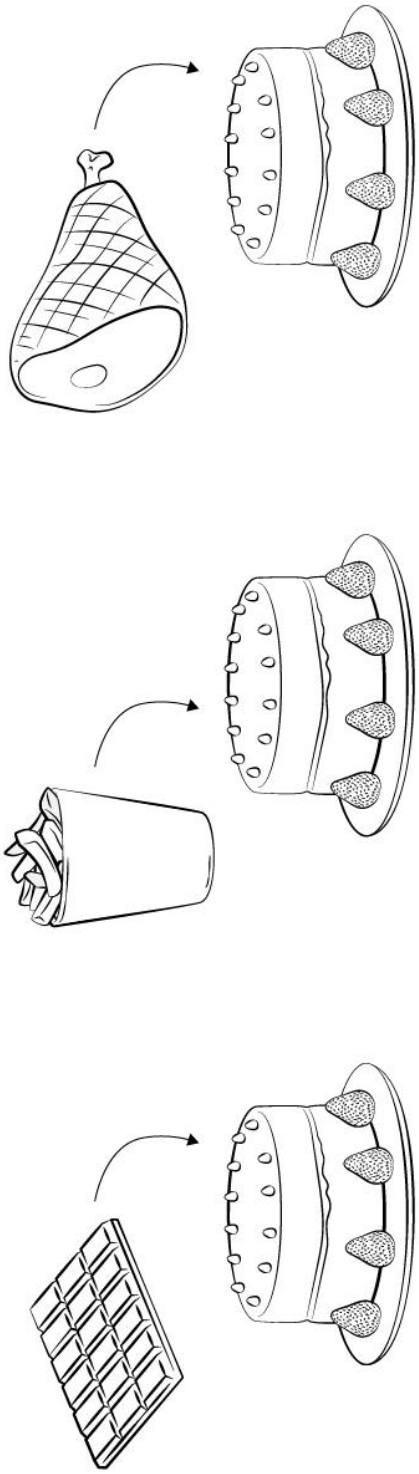




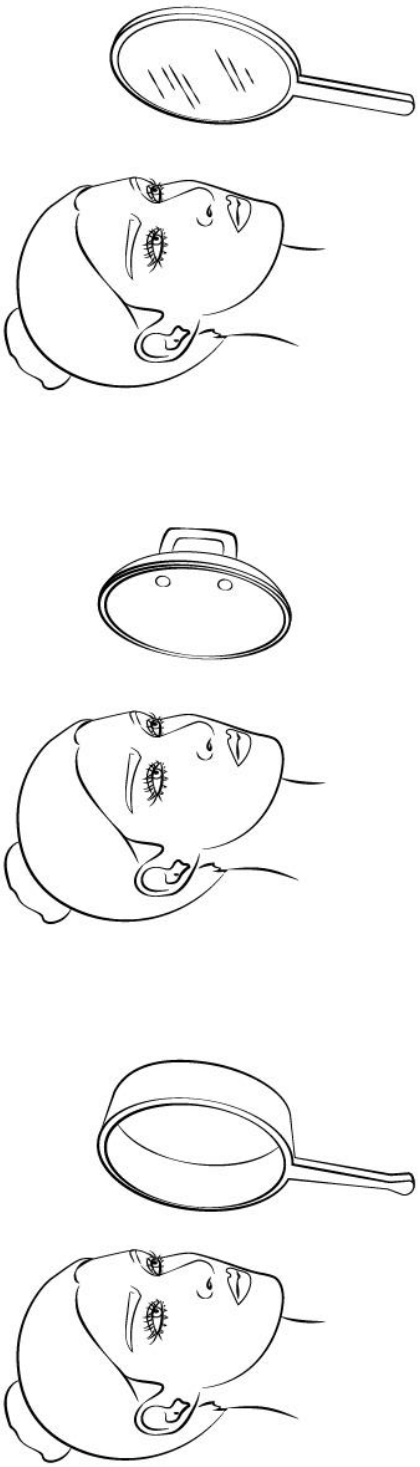
Item 9



Item 10

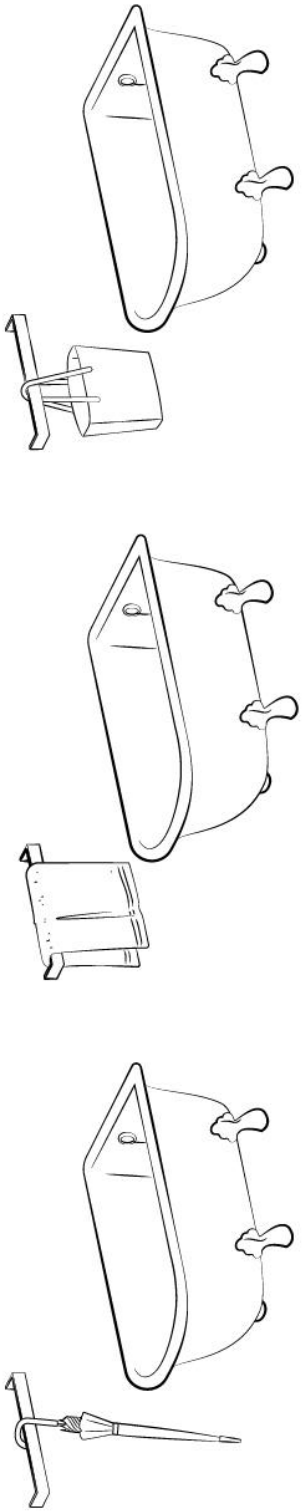


Item 11

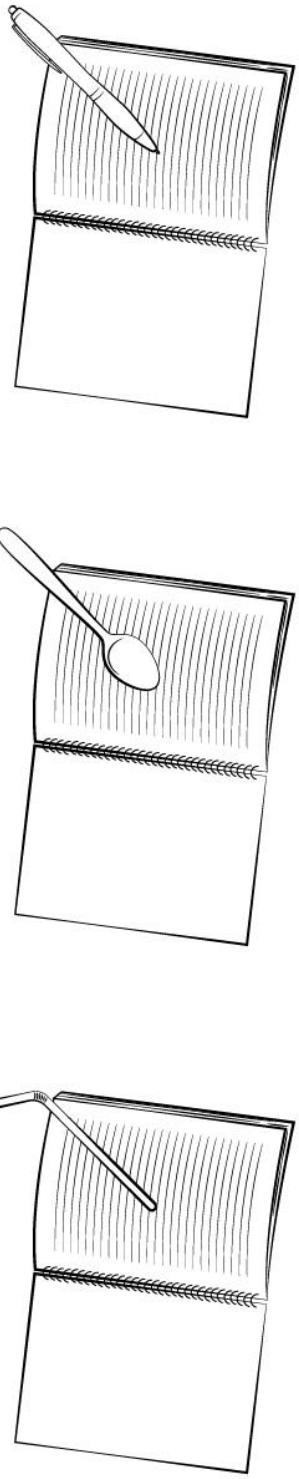


Item 12

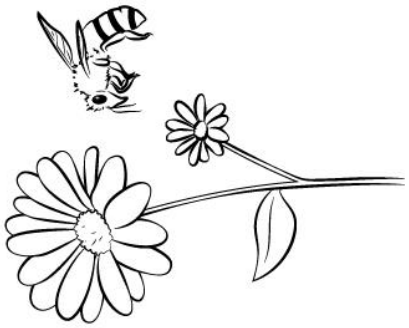
Item 13

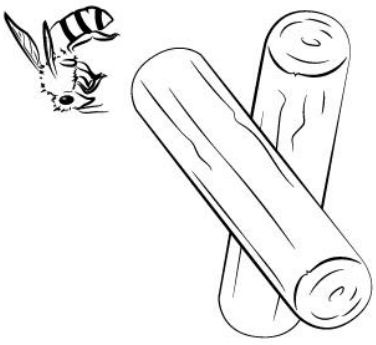


Item 14



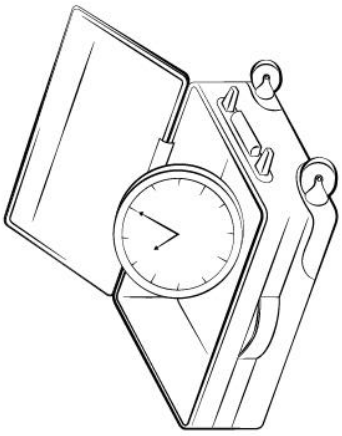
Item 15

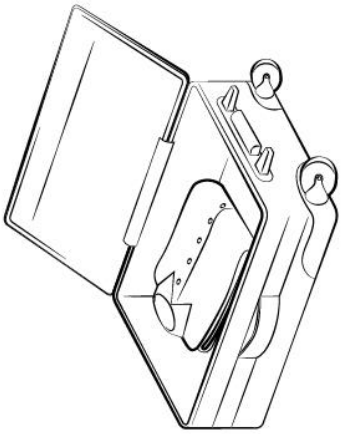


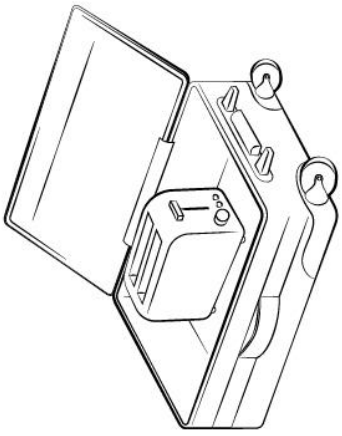


Item 16

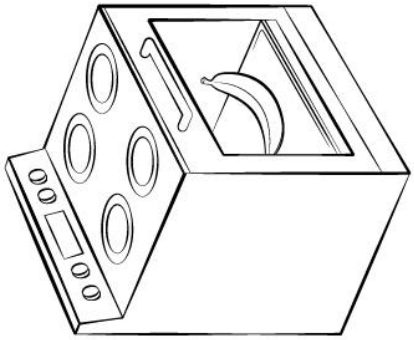
Item 17

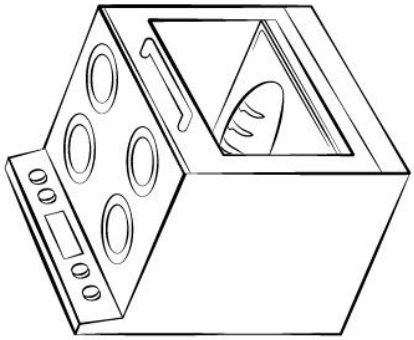


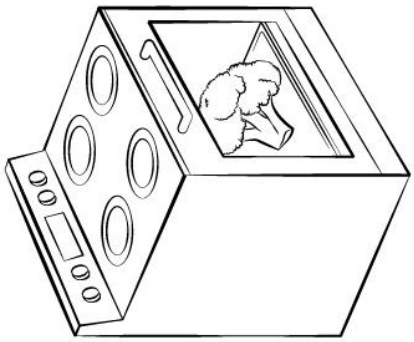




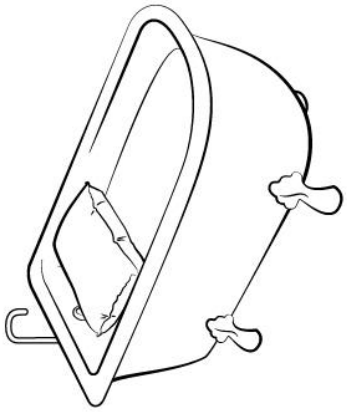
Item 18

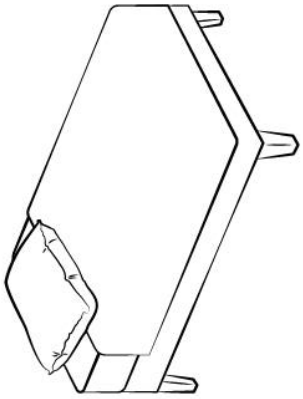


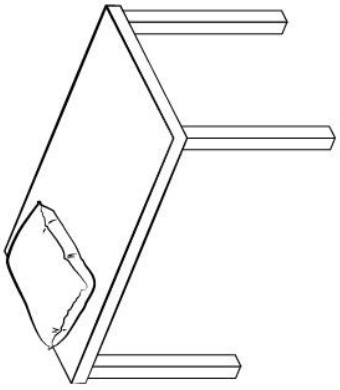




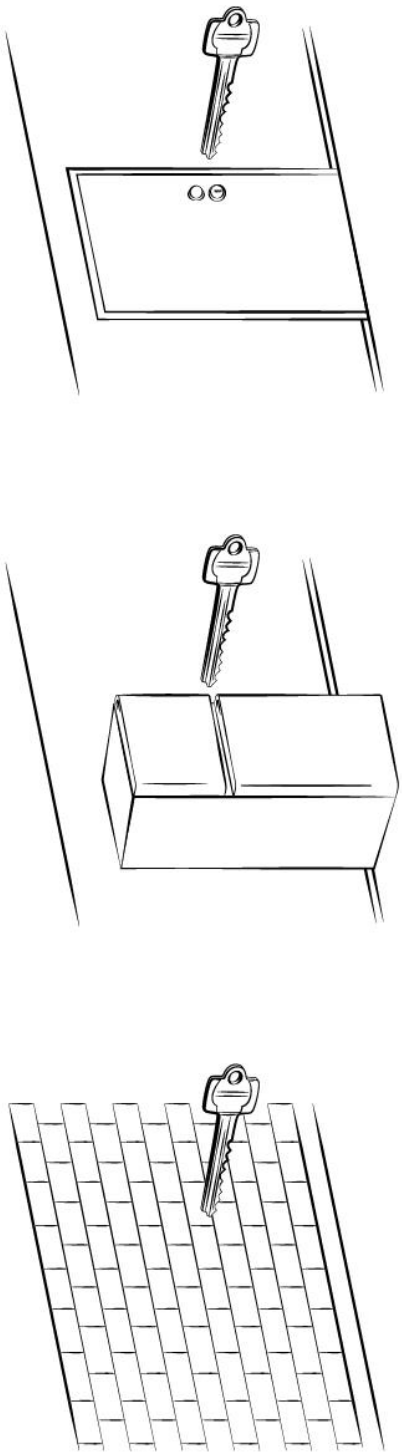
Item 19



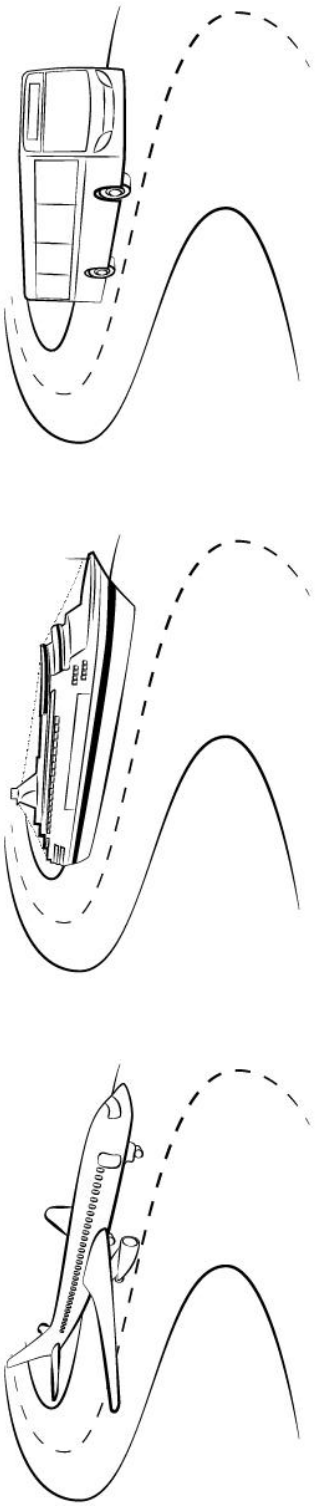




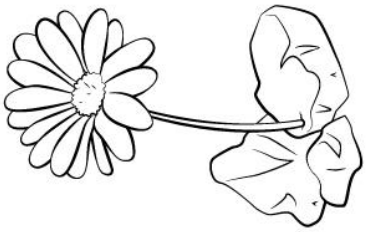
Item 20

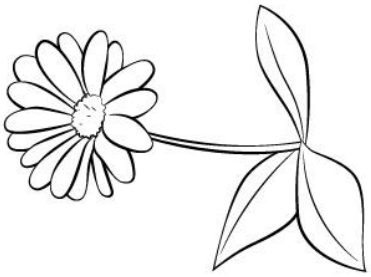


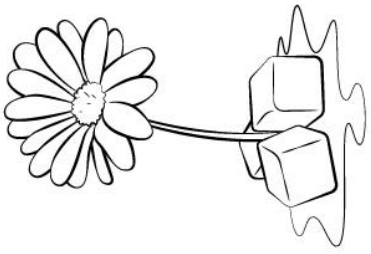
Item 21



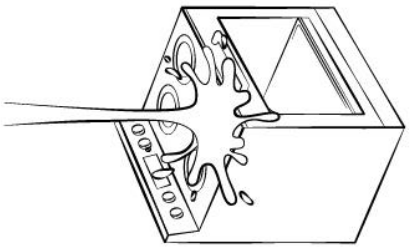
Item 22

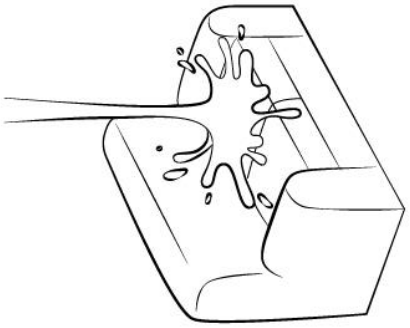


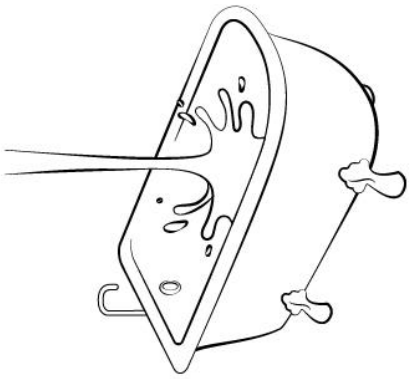




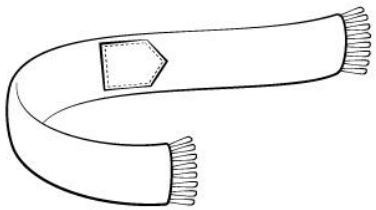
Item 23

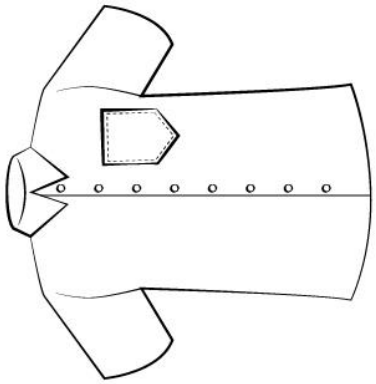


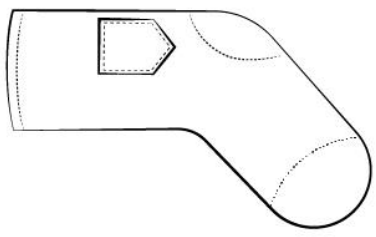




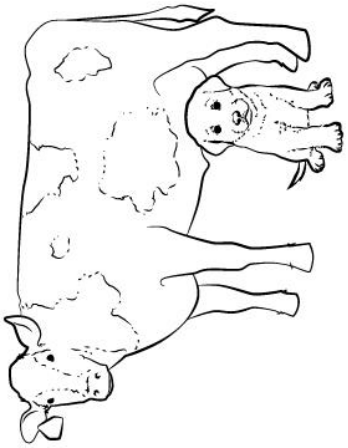
Item 24

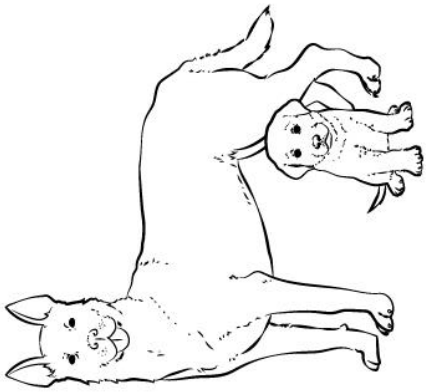




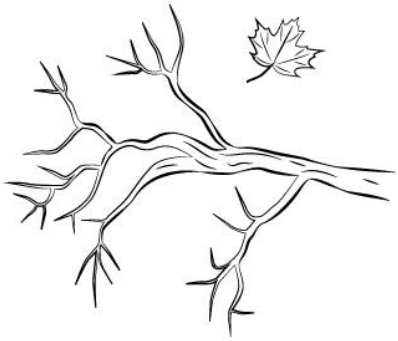


Item 25

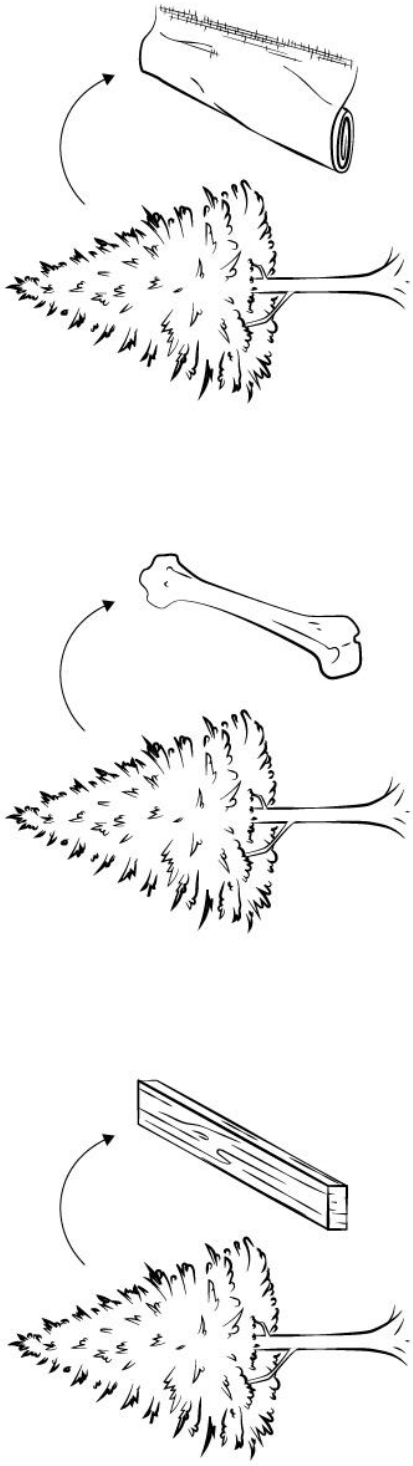




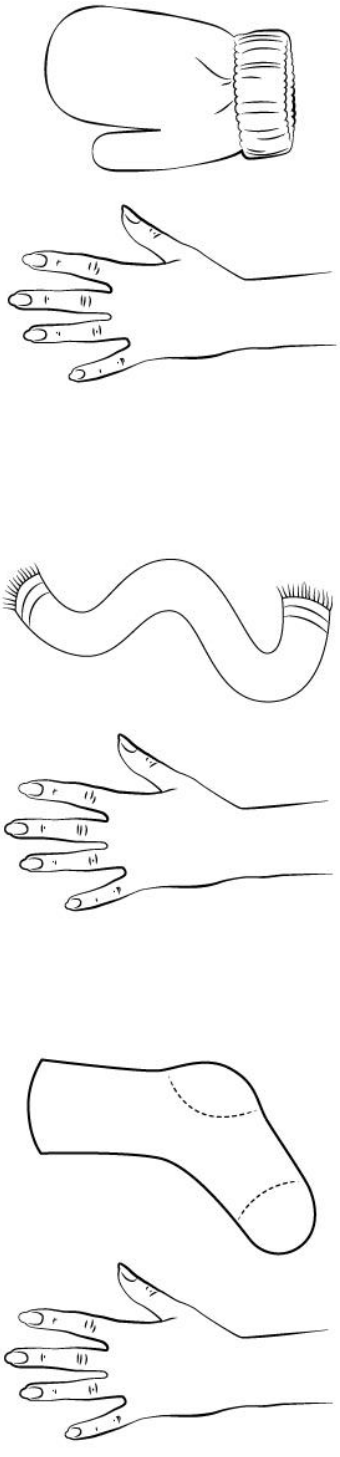
Item 26



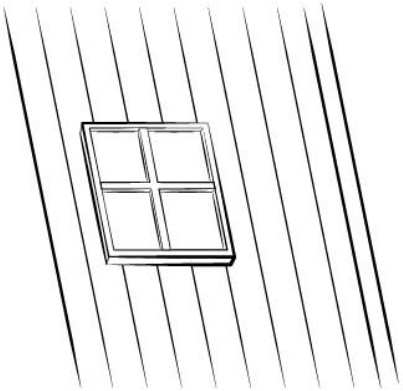
Item 27

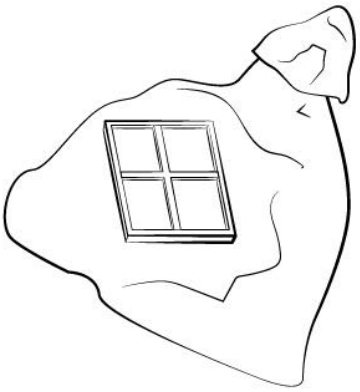


Item 28

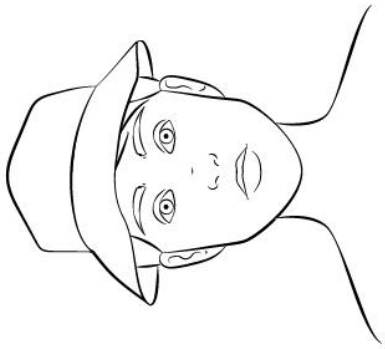


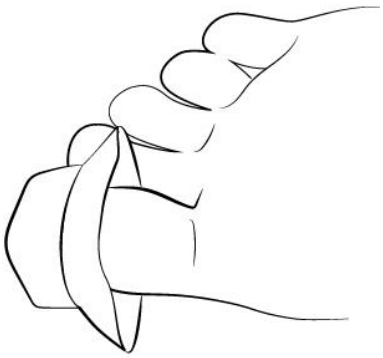
Item 29



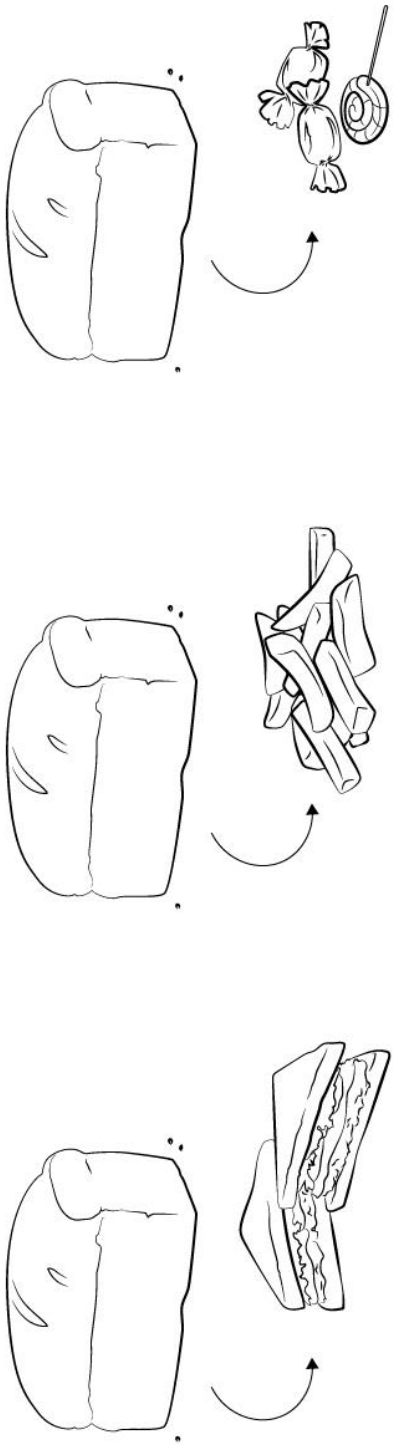


Item 30

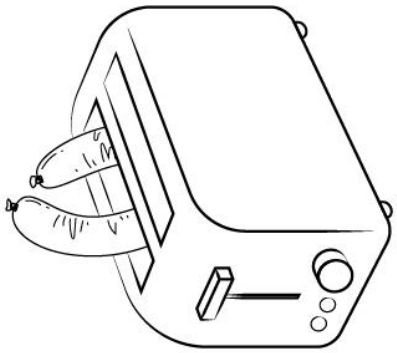


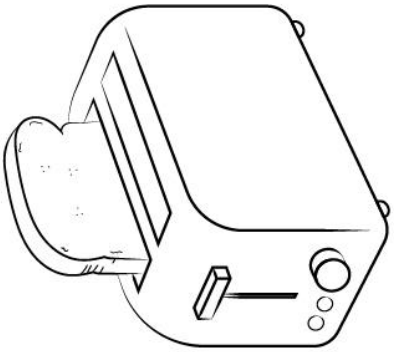


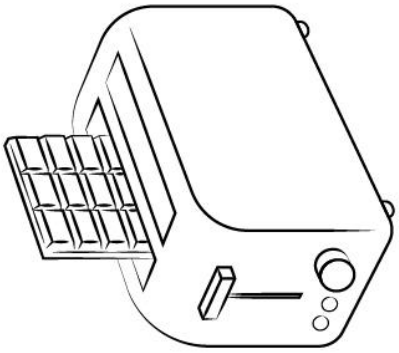
Item 31



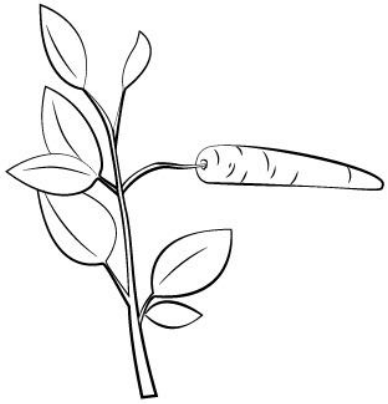
Item 32

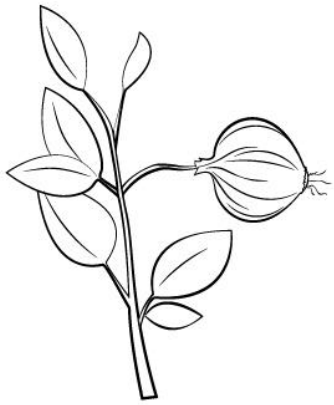


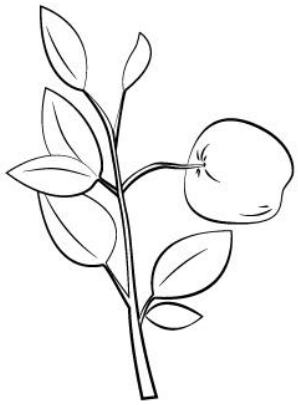




Item 33

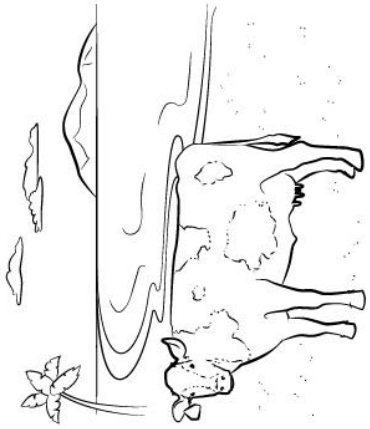


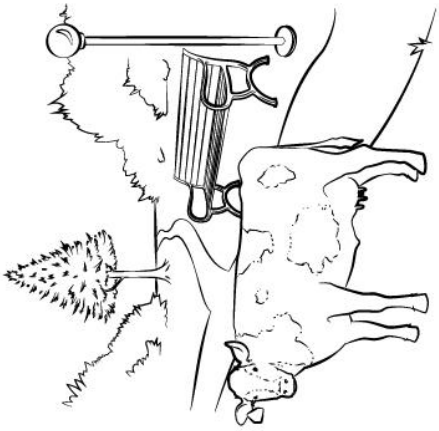


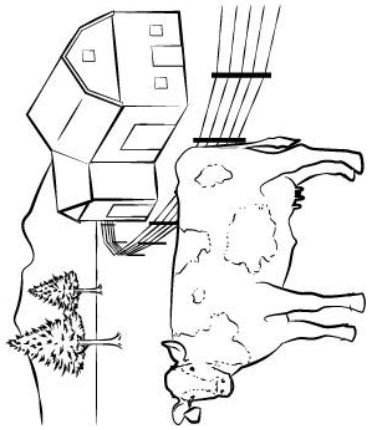


Item 34

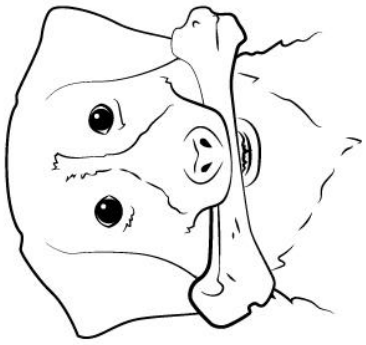
Item 35

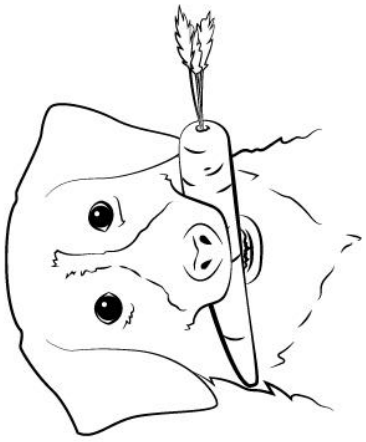


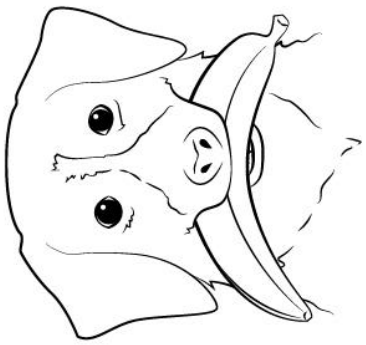




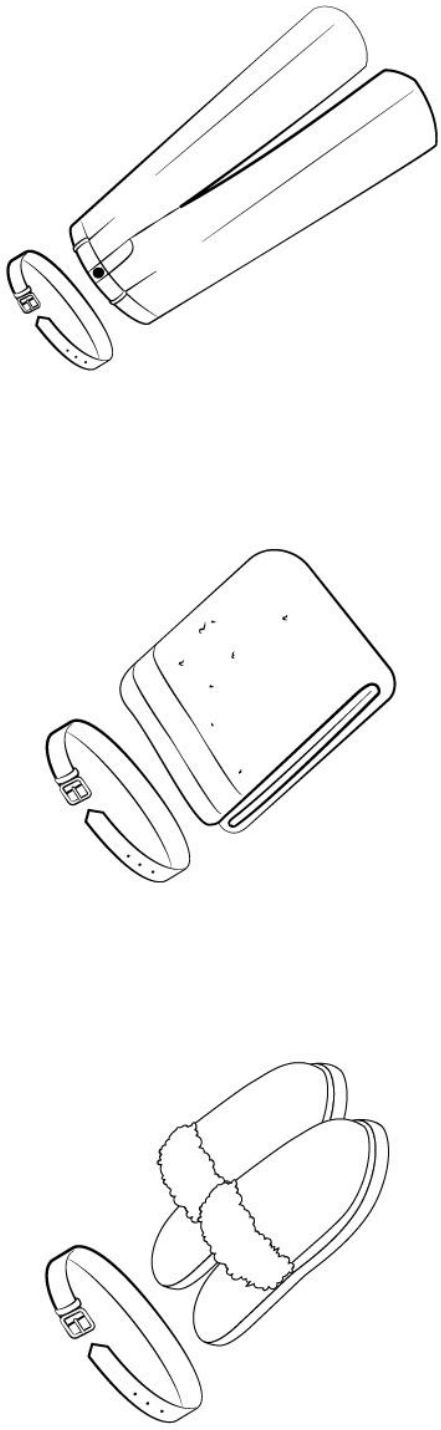
Item 36



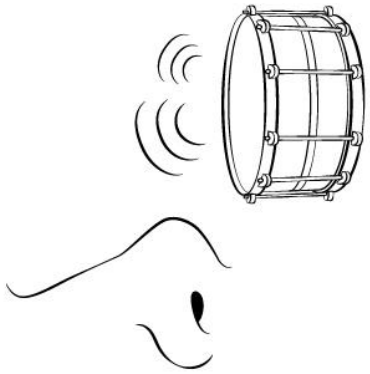


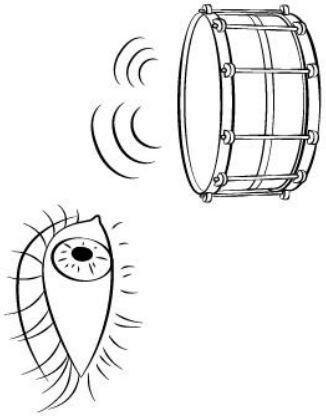


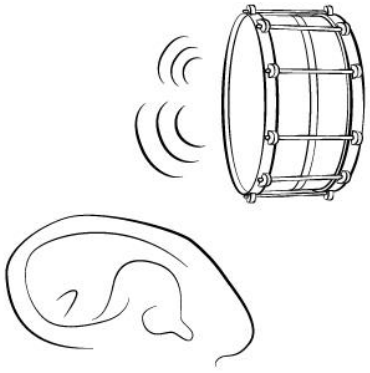
Item 37



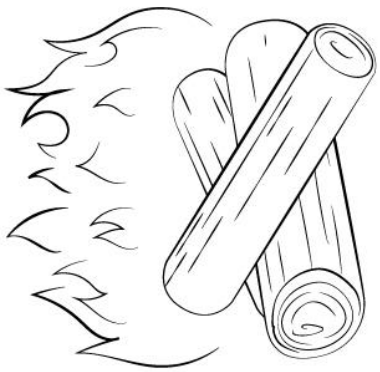
Item 38



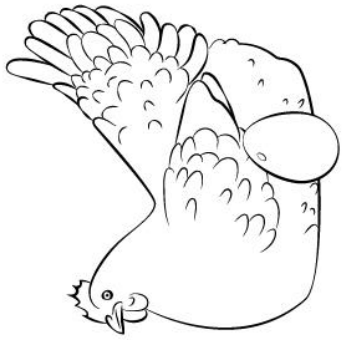


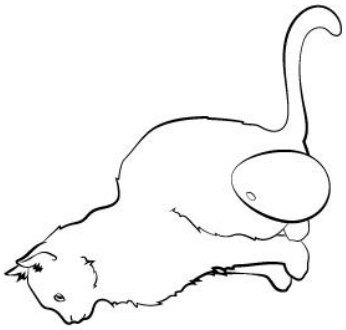


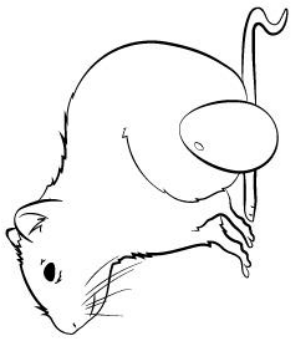
Item 39



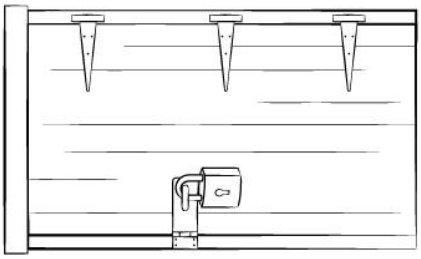
Item 40

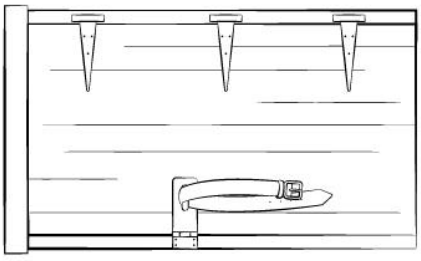


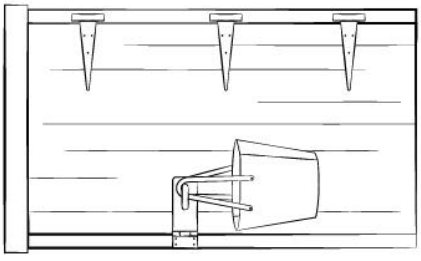




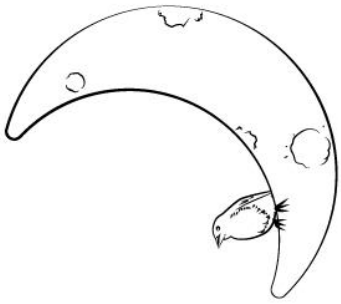
Item 41



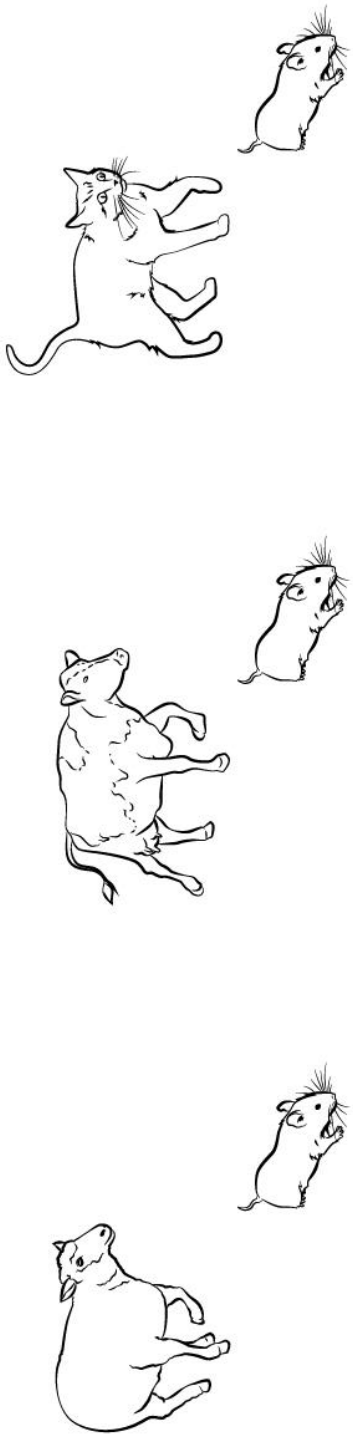




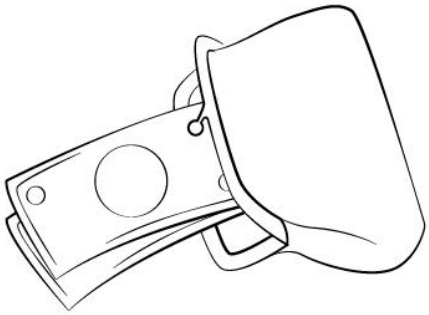
Item 42

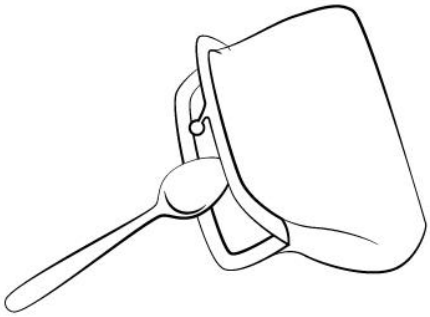


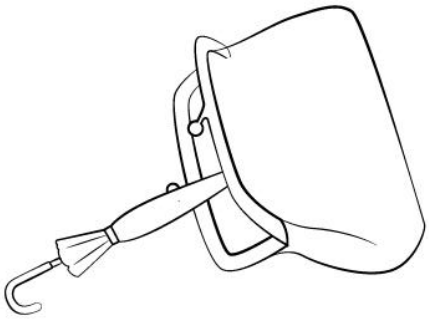
Item 43



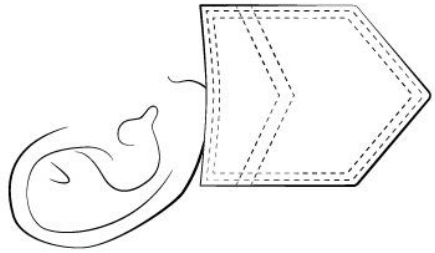
Item 44

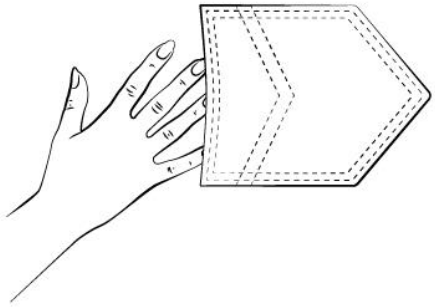


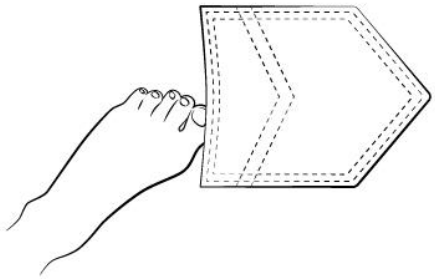




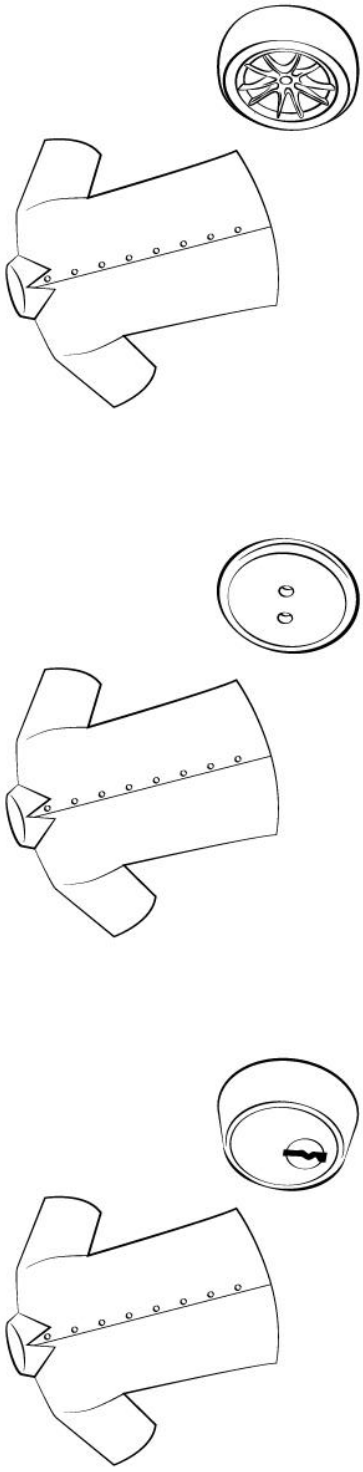
Item 45



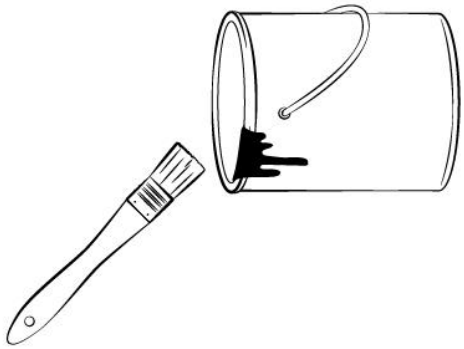


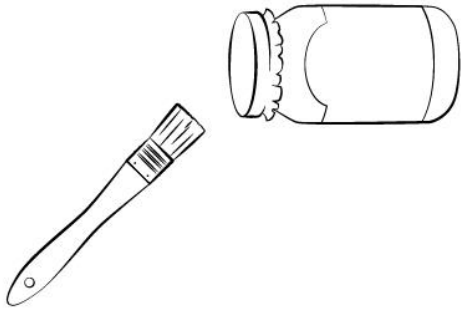


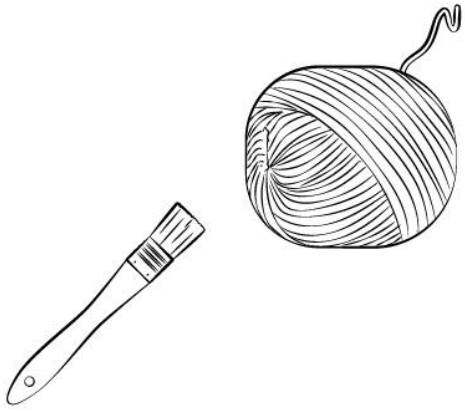
Item 46



Item 47

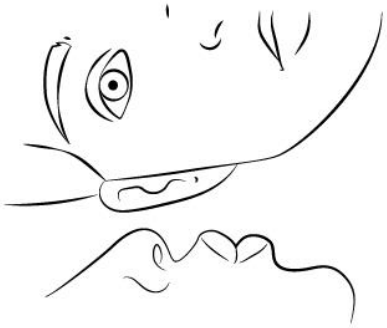


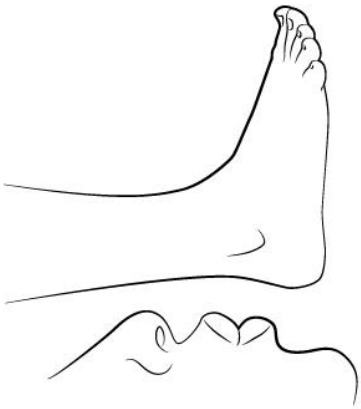




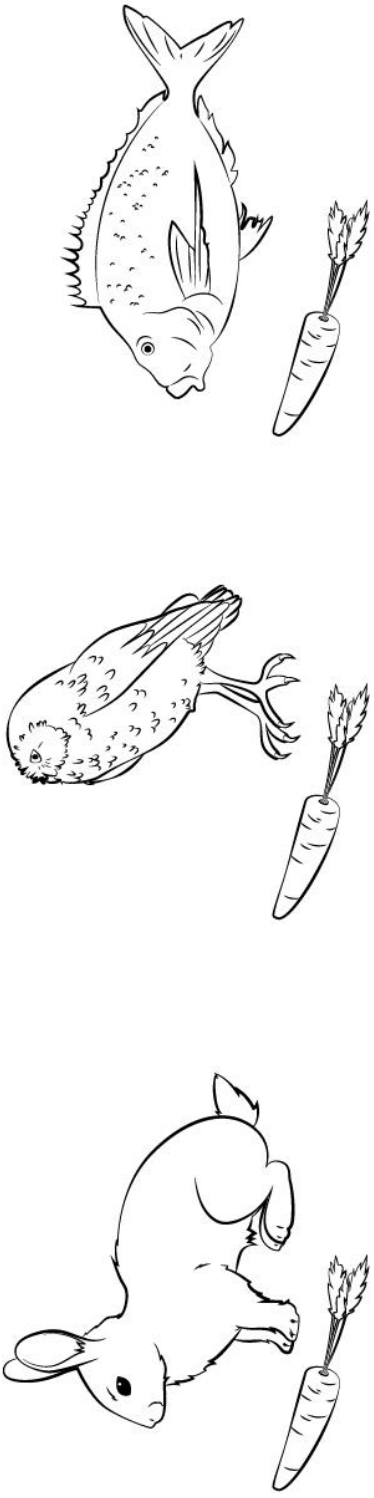
Item 48

Item 49



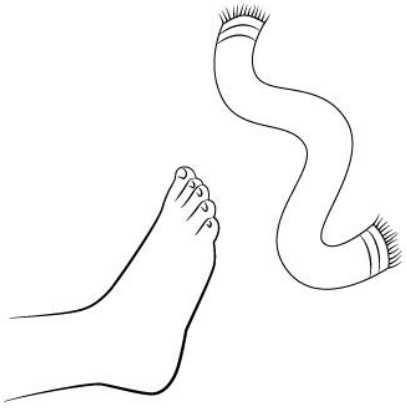


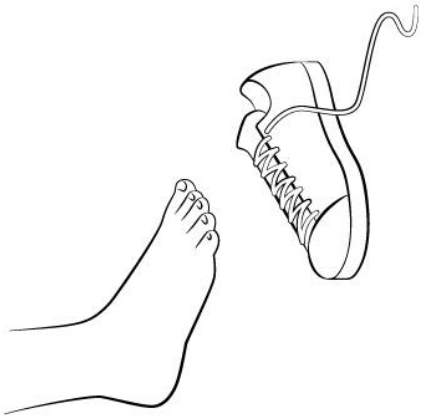
Item 50



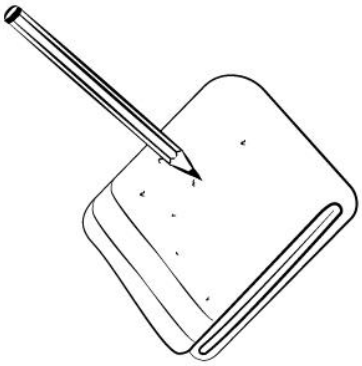
Item 51

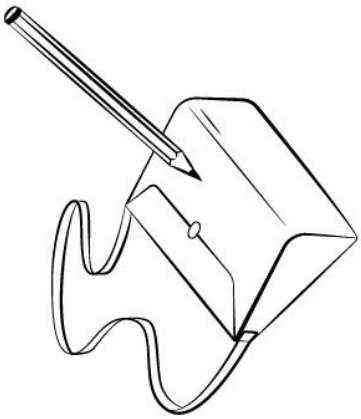
\

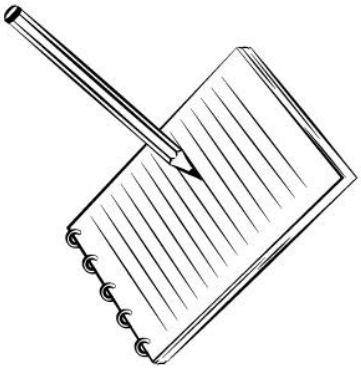




Item 52

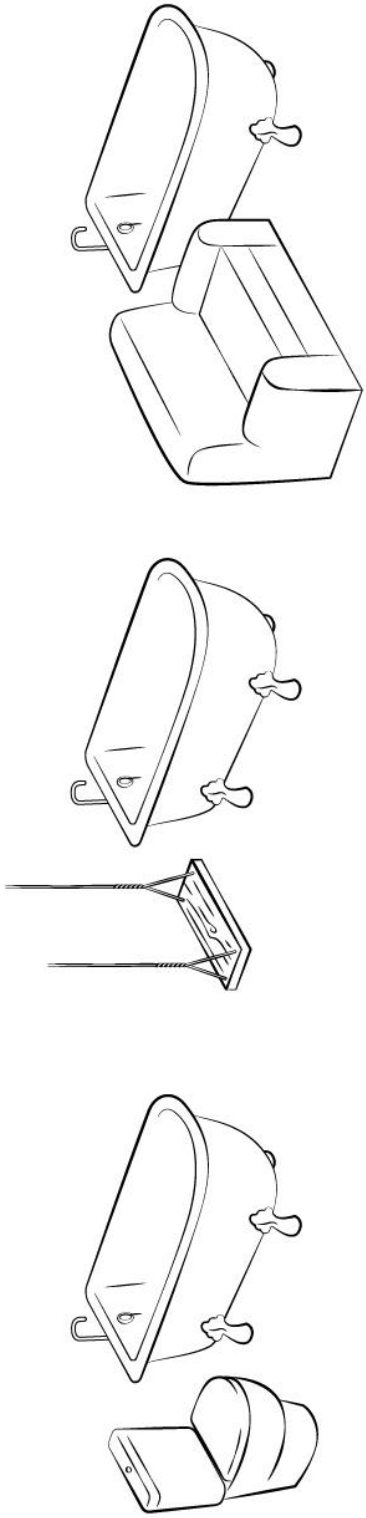




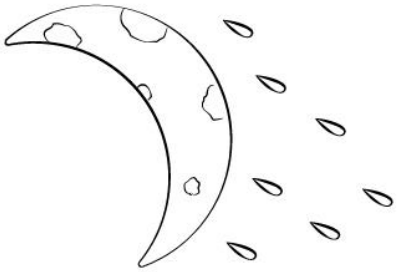


Item 53

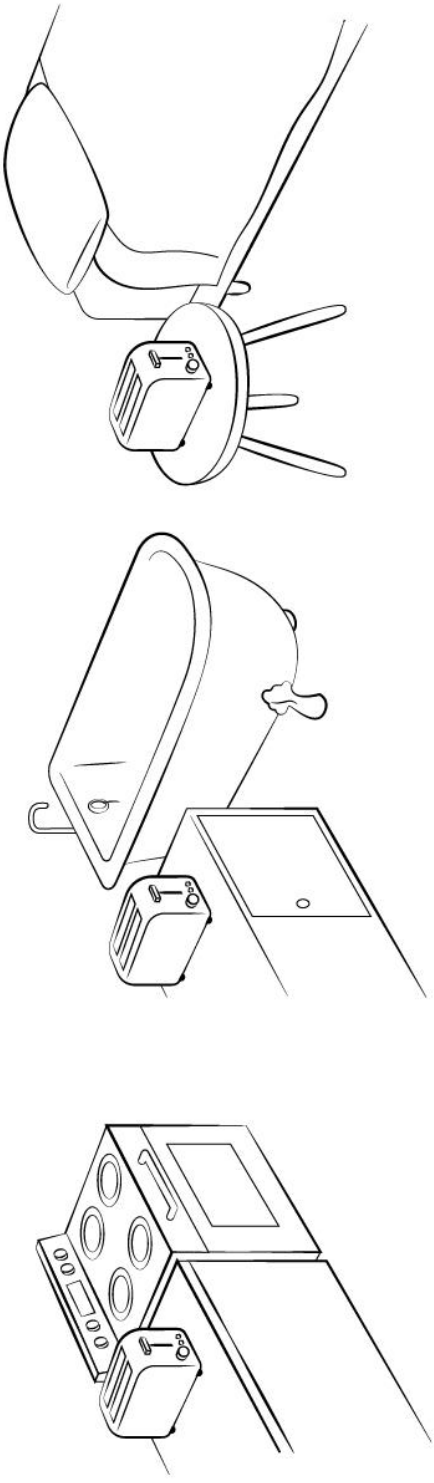
Item 54



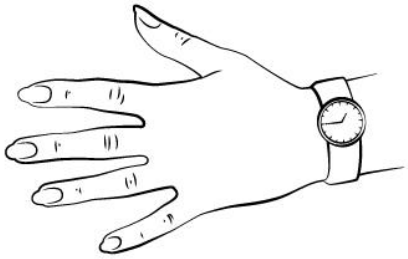
Item 55

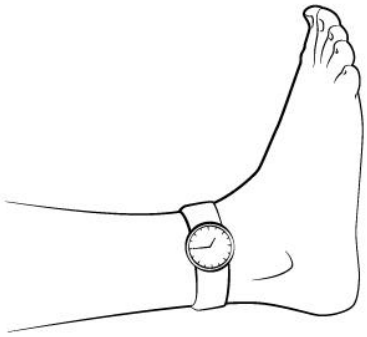


Item 56

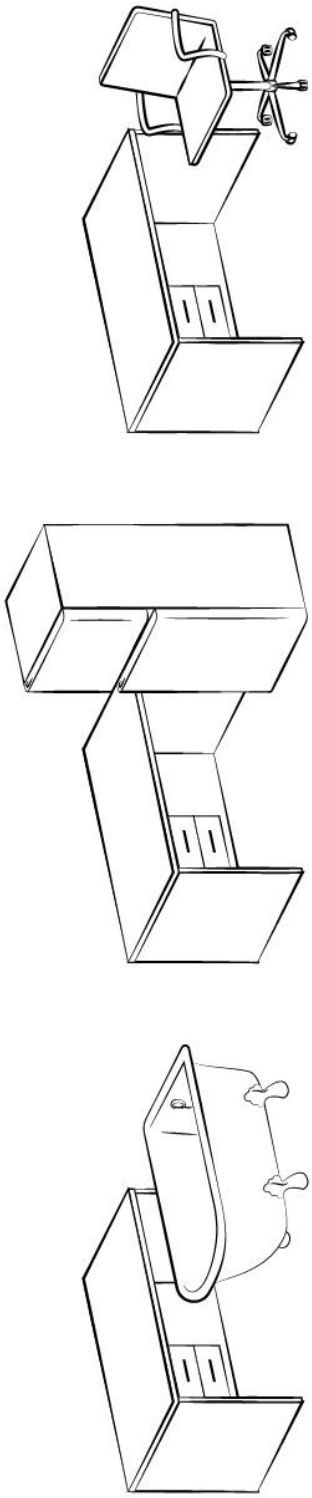


Item 57





Item 58



Appendix B: General Population Validation Ethics

Low Risk Ethics Notification Letter for General Population Study

HoU Review Group

A/Pro Ross Flett

Ethics Notification Number: 4000021031

Title: Mapping the IFOF with a new test of semantic association - general population validation

Thank you for your notification which you have assessed as Low Risk.

Your project has been recorded in our system which is reported in the Annual Report of the Massey University Human Ethics Committee.

The low risk notification for this project is valid for a maximum of three years.

Please note that travel undertaken by students must be approved by the supervisor and the relevant Pro Vice-Chancellor and be in accordance with the Policy and Procedures for Course-Related Student Travel Overseas. In addition, the supervisor must advise the University's Insurance Officer.

A reminder to include the following statement on all public documents:
"This project has been evaluated by peer review and judged to be low risk. Consequently it has not been reviewed by one of the University's Human Ethics Committees. The researcher(s) named in this document are responsible for the ethical conduct of this research.
If you have any concerns about the conduct of this research that you want to raise with someone other than the researcher(s), please contact Professor Craig Johnson, Director (Research Ethics), email humanethics@massey.ac.nz. "

Please note that if a sponsoring organisation, funding authority or a journal in which you wish to publish require evidence of committee approval (with an approval number), you will have to complete the application form again answering yes to the publication question to provide more information to go before one of the University's Human Ethics Committees. You should also note that such an approval can only be provided prior to the commencement of the research.

You are reminded that staff researchers and supervisors are fully responsible for ensuring that the information in the low risk notification has met the requirements and guidelines for submission of a low risk notification.

Yours sincerely

Professor Craig Johnson
Chair, Human Ethics Chairs' Committee and
Director (Research Ethics)

Appendix C: General Population.Validation Survey

Screenshots of the Qualtrics survey used for the general population validation study

Testing Object Association

This research is being conducted by Jo Chapman as part of the requirements for a Doctor of Clinical Psychology degree. Her supervisors are Professor Janet Leathem, Dr Stephen Hill, and Dr Kwok-Keung Leung.

We are developing a new test to be used during treatment for brain tumours. The test is intended to determine if patients have difficulty understanding how objects relate to one another. This will help doctors decide how to best treat tumours safely. Before trying the new test with patients we need to make sure most people can easily answer the questions.

To participate in this research you must:

- be over 18
- never have experienced any significant brain disease or injury, such as:
  - brain tumour
  - stroke
  - traumatic brain injury
  - dementia

This survey should take you around 5-10 minutes to complete. You will be asked some basic information about yourself and then to select the correct picture from each of 71 sets of three pictures, like these:

You have the right to decline to answer any particular question. Completion of the survey implies consent. The data collected is anonymous and will be stored securely. Reports or publications will not include any identifying details.

If you have any questions or wish to discuss this research please e-mail Joanne.Chapman.2@uni.massey.ac.nz or J.M.Leathem@massey.ac.nz or S.R.Hill@massey.ac.nz

This project has been evaluated by peer review and judged to be low risk. Consequently it has not been reviewed by one of the University's Human Ethics Committees. The researcher(s) named in this document are responsible for the ethical conduct of this research. If you have any concerns about the conduct of this research that you want to raise with someone other than the researcher(s), please contact Professor Craig Johnson, Director (Research Ethics), email humanethics@massey.ac.nz

I understand and consent, begin the survey

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

First, please tell us some information about yourself.
This helps us to check whether the new test will be able to used with a wide variety of people.

What is your age?

- ○ 18 - 29
- ○ 30 - 39
- ○ 40 - 49
- ○ 50 - 59
- ○ 60 - 69
- ○ 70 - 79
- ○ 80 or older

What gender do you identify as?

- ○ Male
- ○ Female
- ○ Nonbinary / Genderfluid / Other

What race or ethnicity do you identify with?
Select all that apply.

- ☐ White / European / Caucasian
- ☐ NZ Māori or Pacific Islander
- ☐ Latino / Hispanic
- ☐ East Asian or Asian Indian
- ☐ South Asian or Indian
- ☐ Middle Eastern or North African
- ☐ African
- ☐ African American
- ☐ American Indian or other North American Native
- ☐ Other

Where do you currently live?

What is the highest educational qualification you have received?

- no formal qualification
- a secondary school qualification (high school)
- completed secondary school (college/university entrance or equivalent)
- trade/technical/vocational training
- undergraduate degree (bachelors or equivalent)
- postgraduate degree (masters, doctorate or equivalent)

←

→

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

Next you will see the test questions.

For each question, you will be shown three images and asked which image you think is correct.

For example, which of these is correct?

In this example, the second image is correct because the sun and a cloud can appear together in the sky. However, neither pencils nor fridges are likely to be in the clouds!

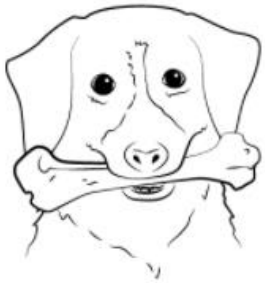
All the test questions are similar. For each question, select which image you think is correct.
If you're not sure, take a guess.
The next question will then be displayed automatically.

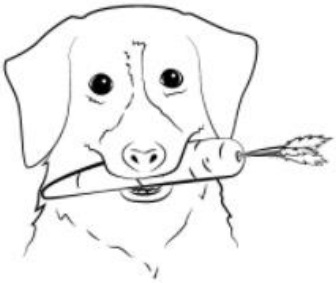
If you have trouble with any of the questions (e.g. telling what the images are or choosing which is correct), please tell us about it at the end of the survey.

← Start test

MASSEY UNIVERSITY
TE KUNENGA KI PUREHUROA
UNIVERSITY OF NEW ZEALAND

Select the correct image.





Thank you, you have completed the test.

Did you have any difficulty with any of the questions?
e.g. telling what the images were or choosing which was correct.

If so, please tell us about it.

←

End survey

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

Testing Object Association

Thank you for taking the time to help with this research.

Your response has been recorded.

If you are interested in updates about how the project is progressing, we will post these on our Facebook page.

Appendix D: Proposed Test Item Issues and Decisions

Summary of issues identified and decision rationale for each proposed test item

Item	Issues identified	Comment themes (where applicable)	Decision	Rationale
1	-		Keep	
2	-		Keep	
3	-		Keep	
4	-		Keep	
5	-		Keep	
6	-		Keep	
7	-		Keep	
8	-		Keep	
9	-		Keep	
10	-		Keep	
11	-		Keep	
12	Comments	Bird identified as seagull	Change image	Not many incorrect answers, but easy to improve based on comments
13	Incorrect/Skipped Comments Cultural effect	Ketchup misidentified as mustard or mayonnaise	Remove	Considered changing 'egg' distractor to a vegetable – too large a change
14	-		Keep	
15	Comments	Dog misidentified as wolf	Change image	Not many incorrect answers, but easy to

				improve based on comments
16	Incorrect/Skipped Age effect		Change image	Change bongo drum to snare drum – older people may have misidentified as a glass/mug/tankard, one comment suggests this
17	-		Keep	
18	-		Keep	
19	-		Keep	
20	-		Keep	
21	Incorrect/Skipped Comments Culture effect Age effect	Difficulty identifying potato	Remove	Clear depiction of potato very difficult
22	Incorrect/Skipped Comments Age effect Culture effect	Difficulty recognising key due to size of image	Change image	Increase size of key – would disproportionately affect older people
23	-		Keep	
24	-		Keep	
25	-		Keep	
26	-		Keep	
27	-		Keep	
28	Comments Cultural effect?	Owls hunt many animals	Remove	Consider changing cat to cow – too large a change
29	-		Keep	
30	-		Keep	
31	-		Keep	
32	-		Keep	
33	-		Keep	
34	-		Keep	
35	-		Keep	
36	-		Keep	
37	-		Keep	
38	-		Keep	
39	-		Keep	
40	Incorrect/Skipped Time Comments	Soup misidentified as cereal or porridge	Remove	Clearer image of soup not possible without colour or other objects
41	-		Keep	
42	-		Keep	

43	-		Keep	
44	-		Keep	
45	-		Keep	
46	Incorrect/Skipped Time Comments Cultural effect	Rice misidentified as grated cheese	Remove	Unable to identify alternative distractor food
47	Incorrect/Skipped Time Comments Cultural effect Age effect		Remove	Clear depiction of salt not possible
48	-		Keep	
49	-		Keep	
50	-		Keep	
51	Incorrect/Skipped Time Comments Age effect	'Brain food' Unhealthy food causing heart disease	Remove	
52	-		Keep	
53	-		Keep	
54	-		Keep	
55	-		Keep	
56	-		Keep	
57	Cultural effect?		Keep	
58	-		Keep	
59	Cultural effect		Remove	Milk containers in different countries esp India
60	Cultural effect		Change image	Move cow closer to sea - background 'bigger'
61	Incorrect/Skipped Time Comments Age effect		Remove	
62	-		Keep	
63	Incorrect/Skipped Comments Cultural effect Age effect	Jam misidentified as various types of preserves	Remove	Unable to identify alternative distractor foods
64	Comments	Kissing interpreted as whispering	Keep	Different interpretations of the image still produce the same answer

65	Incorrect/Skipped Comments Cultural effect Age effect	Paper misidentified as lasagne sheets	Remove	Clear depiction of paper very difficult
66	Incorrect/Skipped Comments Age effect	Photo misidentified as mirror or iPad	Remove	Too many incorrect, framed photo could go on wall with clock
67	-		Keep	
68	-		Keep	
69	Cultural effect Age effect	Watch misidentified as FitBit	Change image	
70	Cultural effect?		Keep	
71	Incorrect/Skipped Time Comments Age effect	Difficulty identifying sugar cubes	Remove	Clear depiction of sugar not possible

Appendix E: Clinical Validation Ethics

Health and Disability Ethics Committee approval for clinical validation study

Health and Disability Ethics Committees
Ministry of Health
133 Molesworth Street
PO Box 5013
Wellington

6011

0800 4
ETHICS
hdecs@health
.govt.nz

30 June 2020

Miss Jo Chapman

Palmerston North 4412

Dear Miss Chapman,

Re:	**Ethics ref:**	**20/NTB/93**
	Study title:	Validating a new test of semantic association for use in mapping the IFOF during awake craniotomy

I am pleased to advise that this application has been *approved* by the Northern B Health and Disability Ethics Committee. This decision was made through the HDEC-Full Review pathway.

Summary of outstanding ethical issues

The main ethical issues considered by the Committee which require addressing by the Researcher are as follows.

The Committee has noted the following – There are three main areas to update as part of PA: **1.** update the study protocol, considering the feedback provided by the Committee. The protocol has been updated with new content re "safety considerations" and digital consent, **2.** Please provide a safety plan addressing the concerns raised by the Committee There doesn't appear to be a separate safety plan uploaded; other than the amendments to the protocol, **3.** Please update the participant information sheet and consent form, taking into account feedback provided by the Committee - re digital consent, the Committee also noted face shot of consent is unusual and suggested an audio recording and written confirmation via email would be more appropriate.

Conditions of HDEC approval

HDEC approval for this study is subject to the following conditions being met prior to the commencement of the study in New Zealand. It is your responsibility, and that of the study's sponsor, to ensure that these conditions are met. No further review by the Northern B Health and Disability Ethics Committee is required.

Standard conditions:

1. Before the study commences at *any* locality in New Zealand, all relevant regulatory approvals must be obtained.

2. Before the study commences at *each given* locality in New Zealand, it must be authorised by that locality in Online Forms. Locality authorisation confirms that the locality is suitable for the safe and effective conduct of the study, and that local research governance issues have been addressed.

After HDEC review

Please refer to the *Standard Operating Procedures for Health and Disability Ethics Committees* (available on www.ethics.health.govt.nz) for HDEC requirements relating to amendments and other post-approval processes.

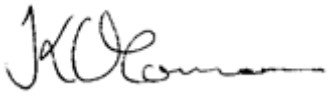
Your next progress report is due by 30 June 2021.

Participant access to ACC

The Northern B Health and Disability Ethics Committee is satisfied that your study is not a clinical trial that is to be conducted principally for the benefit of the manufacturer or distributor of the medicine or item being trialled. Participants injured as a result of treatment received as part of your study may therefore be eligible for publicly-funded compensation through the Accident Compensation Corporation (ACC).

Please don't hesitate to contact the HDEC secretariat for further information. We wish you all the best for your study.

Yours sincerely,



Chairperson
Northern B Health and Disability Ethics Committee

Encl: appendix A: documents submitted
appendix B: statement of compliance and list of members

Appendix A
Documents submitted

Document	*Version*	*Date*
CV for CI	1.0	16 March 2020
Protocol: Updated with tracked changes for provisional approval response	1.1	02 April 2020
PIS/CF: Information for current glioma patients	0.7	14 April 2020
PIS/CF: Information for discharged glioma patients	0.5	14 April 2020

PIS/CF: Information for semantic impairment patients	0.5	14 April 2020
PIS/CF: Information for semantic impairment patients in Easy Read format	0.6	14 April 2020
Information for Speech Language Therapists	0.4	06 April 2020
PIS/CF: Exported content of online consent form	0.2	14 April 2020
Evidence of scientific review: Peer review form completed by chair of comfirmation panel.	1.0	15 April 2020
Application		16 April 2020
Covering Letter: Covering letter for provisional approval response	1.0	05 June 2020
Response to Request for Further Information		

Appendix B
Statement of compliance and list of members

Statement of compliance

The Northern B Health and Disability Ethics Committee:

- is constituted in accordance with its Terms of Reference
- operates in accordance with the *Standard Operating Procedures for Health and Disability Ethics Committees*, and with the principles of international good clinical practice (GCP)
- is approved by the Health Research Council of New Zealand's Ethics Committee for the purposes of section 25(1)(c) of the Health Research Council Act 1990
- is registered (number 00008715) with the US Department of Health and Human Services' Office for Human Research Protection (OHRP).

List of members

Name	*Category*	*Appointed*	*Term Expires*
Mr John Hancock	Lay (the law)	14/12/2015	14/12/2018
Dr Nora Lynch	Non-lay (health/disability service provision)	24/07/2015	24/07/2022
Miss Tangihaere Macfarlane	Lay (consumer/community perspectives)	20/05/2017	20/05/2020
Mrs Kate O'Connor	Lay (ethical/moral reasoning)	14/12/2015	14/12/2018
Mrs Stephanie Pollard	Non-lay (intervention studies)	01/07/2015	01/07/2018
Mrs Leesa Russell	Non-lay (intervention studies), Nonlay (observational studies)	14/12/2015	14/12/2018
Ms Susan Sherrard	Lay (consumer/community perspectives)	19/03/2019	19/03/2022
Mrs Jane Wylie	Non-lay (intervention studies)	20/05/2017	20/05/2020

Unless members resign, vacate or are removed from their office, every member of HDEC shall continue in office until their successor comes into office (HDEC Terms of Reference)

http://www.ethics.health.govt.nz

Appendix F: Clinical Validation Recruitment

This appendix contains the following documents, related to recruitment for the clinical validation study (described in chapter 6):

1. Information Sheets
    - Participant with Glioma
    - Participant with Semantic Impairment
    - Easy Read
    - Speech Language Therapist
2. Recruitment Quick Reference guide for Neurology department staff

TESTING OBJECT ASSOCIATION

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

PARTICIPANT INFORMATION SHEET

Ethics committee reference: 20/NTB/93
Lead investigator: Jo Chapman

You are invited to take part in a study of a new test of brain function.

This information will help you decide if you'd like to take part. It sets out why we are doing the study, what taking part would involve, possible benefits and risks, and what will happen after the study ends.

Whether or not you take part is your choice. You do not have to decide today. Before you decide you may want to talk about the study with other people, such as family, whānau, friends, or healthcare providers.

If you do want to take part in this study, you will fill in a consent form online when we meet. An image of the consent form is shown on the last page of this document. You can pull out of the study at any time if you change your mind.

If you don't want to take part, you don't have to give a reason, and it won't affect the care you receive.

This document is 263 pages long, including the consent form. Please make sure you have read and understood all the pages.

When your doctor gives you this document, please choose whether:

☐ Your doctor will send your contact details to the study team. Someone from the team will then contact you to talk about whether you would like to take part.

OR

☐ You will contact the study team yourself if you want to take part. You can also contact us if you are not sure about taking part or have questions. The person to contact is:

Jo Chapman [email]
[phone number]

WHY WE ARE DOING THIS STUDY

This study aims to find out whether a new test of brain function works better than the tests currently used during brain tumour treatment. These tests are used to help surgeons know which parts of the brain are working well and which are not.

Before we can use the new test in treatment, we need to know whether it works well for people with brain tumours. We can also learn which parts of the brain affect the test by comparing people with tumours in different places.

This study has been approved by the New Zealand Health and Disability Ethics Committee.

WHAT TAKING PART WILL INVOLVE

If you would like to take part in this study, we will meet with you via video conference. The meeting will take about 45 minutes and be at a time that suits you. You can have a family member, friend or other support person with you if you wish.

We will use Zoom for the meeting. Zoom is free and simple to use. We will send you instructions for the meeting and can help with setup if needed. You will need:

- a computer with a video camera and microphone
- a stable internet connection
- a private and quiet place where you will not be interrupted

When we meet, we will first talk through this information sheet and answer any questions you have about the study. We may ask you some questions to check we have explained the study correctly. If you then choose take part, we will fill in the consent form together online. We will also save a picture of you and the completed form as evidence of your identity.

We will then ask about your age, gender, ethnicity, and education. This information will help us understand how the test works with different people.

You will do 3 tests during the meeting: the new test and 2 current tests. During the tests, we will ask you to:

- look at pictures and choose which are correct or go together
- think of words and say them as fast as you can

At the end of the meeting, we will also talk about how doing the tests felt and you can ask us any further questions.

After meeting with you, we will contact the doctor who referred you to this study and ask:

- about your tumour
    - where it is
    - when it started
    - what grade it is

- about your Diffusion Tensor Imaging (DTI) scan(s)
    - how your tumour affects one particular part of your brain
- about your treatment
    - whether you have had surgery or radiation treatment
    - when this treatment was

We will not ask for any other information, see your scans or medical records, or tell your doctor the results of the tests you did with us.

POSSIBLE BENEFITS AND RISKS OF THIS STUDY

Your medical care:

Taking part in this study will not change anything about your medical care.

Your experience:

We will do our best to make meeting with us a positive experience. You may feel upset or worried if you find the tests difficult. We can talk about this if it happens. We may be able to explain why you had trouble or share others' experiences.

You may also enjoy doing the tests and/or feel pleased to have taken part in a study that could help others with tumours in the future.

Your privacy:

We will not ask for any medical information from your doctor until we have met with you and you have completed the consent form. We will keep your medical information and test results separate from your name and identifying information. However, this information is not entirely anonymous because the details of your tumour may be unique. The information will be kept in a secure online database and only the study team will be able to see it. You can ask to see or correct the information we have about you at any time.

When we share the results of the study, we will compare groups of similar people. We will not include any information about individuals. Other studies may use this group information in future but will not have access to any individual details.

Video conferencing:

The software we use is secure, but privacy online cannot be guaranteed. It may still be possible for others to listen to our meeting via the internet. Others could also overhear our conversation in person, so it is important to use a quiet private place for the meeting. We will use a private room so people cannot overhear at our end. We will not record the meeting.

Video conferencing will use some internet bandwidth and data.

Funding:

Massey University is funding this study. You do not need to pay anything to take part. We will send you a $30 grocery voucher to thank you for meeting with us.

WHAT WILL HAPPEN AFTER THE STUDY

If you wish, we will send you a summary of the results of the study at the end of 2021.

We will share the results of the study in scientific journals and in a Massey University Doctor of Clinical Psychology thesis.

Massey University will keep the consent forms and information from this study in a secure place for 10 years, then destroy them. The individual information will not be used for any other study.

WHO TO CONTACT ABOUT THIS STUDY

Jo Chapman is running this study as part of her Doctor of Clinical Psychology degree at Massey University. Her supervisors are Professor Janet Leathem and Dr Stephen Hill from Massey University and Dr Kwok-Keung Leung.

You can contact us if you have any questions, concerns or complaints about the study at any stage:

Jo Chapman	[email]
	[phone number]
Professor Janet Leathem	[email]
Dr Stephen Hill	[email]
Dr Kwok-Keung Leung	[email]

For Māori cultural support, you can contact a Kaimatai Hinengaro Matua: Māori:

Dr Simon Bennett	[email]

If you want to talk to someone who is not involved with the study, you can contact an independent health and disability advocate on:

Phone:	0800 555 050
Email:	advocacy@advocacy.org.nz
Website:	https://www.advocacy.org.nz/

You can also contact the health and disability ethics committee (HDEC) that approved this study on:

Phone:	0800 4 ETHICS
Email:	hdecs@moh.govt.nz

TESTING OBJECT ASSOCIATION

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

EXAMPLE CONSENT FORM

Please tick to indicate you consent to the following:

- I have read, or have had read to me, and I understand the Participant Information Sheet. I have a copy of the information sheet and this form. ☐
- I have been given enough time to consider whether to take part in this study. I have had the opportunity to use a support person to help me ask questions and understand the study. ☐
- I am satisfied with the answers I have been given about the study. I know who to contact if I have any further questions. ☐
- I understand that taking part in this study is my choice and that I may withdraw from the study at any time without affecting my medical care. ☐
- I consent to the study team collecting and processing my information, including specific information from my doctor or therapist about my health. ☐
- I understand that my information is confidential and that no material which could identify me will be used in any reports on this study. ☐
- I understand the video conferencing software is secure, but privacy online cannot be guaranteed. ☐
- I consent to the researcher recording an image of me and this form as evidence of my consent. ☐

I wish to receive a summary of the results from the study.

- Yes ○
- No ○

If yes, please send the summary of results to this email address:

Please send my $30 supermarket voucher to this address:

I consent to take part in this study.

Participant's name:

Signature:

TESTING OBJECT ASSOCIATION

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

PARTICIPANT INFORMATION SHEET

Ethics committee reference: 20/NTB/93
Lead investigator: Jo Chapman

You are invited to take part in a study of a new test of brain function.

This information will help you decide if you'd like to take part. It sets out why we are doing the study, what taking part would involve, possible benefits and risks, and what will happen after the study ends.

Whether or not you take part is your choice. You do not have to decide today. Before you decide you may want to talk about the study with other people, such as family, whānau, friends, or healthcare providers.

If you do want to take part in this study, you will fill in a consent form online when we meet. An image of the consent form is shown on the last page of this document. You can pull out of the study at any time if you change your mind.

If you don't want to take part, you don't have to give a reason, and it won't affect the care you receive.

This document is 263 pages long, including the consent form. Please make sure you have read and understood all the pages.

When your Speech Language Therapist gives you this document, you can choose whether:

☐ Your therapist will send your contact details to the study team. Someone from the team will then contact you to talk about whether you would like to take part.

OR

☐ You will contact the study team yourself if you want to take part. You can also contact us if you are not sure about taking part or have questions. The person to contact is:

Jo Chapman [email]
[phone number]

WHY WE ARE DOING THIS STUDY

This study aims to find out whether a new test of brain function works better than the tests currently used during brain tumour treatment. These tests are used to help surgeons know which parts of the brain are working well and which are not.

People with brain conditions like yours can help us understand how the test works because you have difficulty with only one type of thinking. If you find the test hard, we will know the test measures this difficult type of thinking. If you find the test easy, we will know it measures a different type of thinking.

This study has been approved by the New Zealand Health and Disability Ethics Committee.

WHAT TAKING PART WILL INVOLVE

If you would like to take part in this study, we will meet with you via video conference. The meeting will take about 45 minutes and be at a time that suits you. You can have a family member, friend or other support person with you if you wish.

We will use Zoom for the meeting. Zoom is free and simple to use. We will send you instructions for the meeting and can help with setup if needed. You will need:

- a computer with a video camera and microphone
- a stable internet connection
- a private and quiet place where you will not be interrupted

When we meet, we will first talk through this information sheet and answer any questions you have about the study. We may ask you some questions to check we have explained the study correctly. If you then choose take part, we will fill in the consent form together online. We will also save a picture of you and the completed form as evidence of your identity.

We will then ask about your age, gender, ethnicity, and education. This information will help us understand how the test works with different people.

You will do 3 tests during the meeting: the new test and 2 current tests. During the tests, we will ask you to:

- look at pictures and choose which are correct or go together
- think of words and say them as fast as you can

At the end of the meeting, we will also talk about how doing the tests felt and you can ask us any further questions.

After meeting with you, we will contact your Speech Language Therapist to ask which condition you have. We will only ask about the condition that affects your thinking. We will not ask about other conditions, see your medical records, or tell your therapist the results of the tests you did with us.

POSSIBLE BENEFITS AND RISKS OF THIS STUDY

Your medical care:

Taking part in this study will not change anything about your medical care.

Your experience:

We will do our best to make meeting with us a positive experience. You may feel upset or worried if you find the tests difficult. We can talk about this if it happens. We may be able to explain why you had trouble or share others' experiences.

You may also enjoy doing the tests and/or feel pleased to have taken part in a study that could help others.

Your privacy:

We will not ask for any medical information from your Speech Language Therapist until we have met with you and you have completed the consent form. We will keep your medical information and test results separate from your name and identifying information. However, this information is not entirely anonymous because the details of your condition may be unique. The information will be kept in a secure online database and only the study team will be able to see it. You can ask to see or correct the information we have about you at any time.

When we share the results of the study, we will compare groups of similar people. We will not include any information about individuals. Other studies may use this group information in future but will not have access to any individual details.

Video conferencing:

The software we use is secure, but privacy online cannot be guaranteed. It may still be possible for others to listen to our meeting via the internet. Others could also overhear our conversation in person, so it is important to use a quiet private place for the meeting. We will use a private room so people cannot overhear at our end. We will not record the meeting.

Video conferencing will use some internet bandwidth and data.

Funding:

Massey University is funding this study. You do not need to pay anything to take part. We will send you a $30 grocery voucher to thank you for meeting with us. We also send a grocery voucher to your Speech Language Therapist to thank them for their help in contacting you.

WHAT WILL HAPPEN AFTER THE STUDY

If you wish, we will send you a summary of the results of the study at the end of 2021.

We will share the results of the study in scientific journals and in a Massey University Doctor of Clinical Psychology thesis.

Massey University will keep the consent forms and information from this study in a secure place for 10 years, then destroy them. The individual information will not be used for any other study.

WHO TO CONTACT ABOUT THIS STUDY

Jo Chapman is running this study as part of her Doctor of Clinical Psychology degree at Massey University. Her supervisors are Professor Janet Leathem and Dr Stephen Hill from Massey University and Dr Kwok-Keung Leung.

You can contact us if you have any questions, concerns or complaints about the study at any stage:

Jo Chapman	[email]
	[phone number]
Professor Janet Leathem	[email]
Dr Stephen Hill	[email]
Dr Kwok-Keung Leung	[email]

For Māori cultural support, you can contact a Kaimatai Hinengaro Matua: Māori:

Dr Simon Bennett	[email]

If you want to talk to someone who isn't involved with the study, you can contact an independent health and disability advocate on:

Phone:	0800 555 050
Email:	advocacy@advocacy.org.nz
Website:	https://www.advocacy.org.nz/

You can also contact the health and disability ethics committee (HDEC) that approved this study on:

Phone:	0800 4 ETHICS
Email:	hdecs@moh.govt.nz

TESTING OBJECT ASSOCIATION

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

EXAMPLE CONSENT FORM

Please tick to indicate you consent to the following:

- [ ] I have read, or have had read to me, and I understand the Participant Information Sheet. I have a copy of the information sheet and this form.
- [ ] I have been given enough time to consider whether to take part in this study. I have had the opportunity to use a support person to help me ask questions and understand the study.
- [ ] I am satisfied with the answers I have been given about the study. I know who to contact if I have any further questions.
- [ ] I understand that taking part in this study is my choice and that I may withdraw from the study at any time without affecting my medical care.
- [ ] I consent to the study team collecting and processing my information, including specific information from my doctor or therapist about my health.
- [ ] I understand that my information is confidential and that no material which could identify me will be used in any reports on this study.
- [ ] I understand the video conferencing software is secure, but privacy online cannot be guaranteed.
- [ ] I consent to the researcher recording an image of me and this form as evidence of my consent.

I wish to receive a summary of the results from the study.

- ○ Yes
- ○ No

If yes, please send the summary of results to this email address:

Please send my $30 supermarket voucher to this address:

I consent to take part in this study.

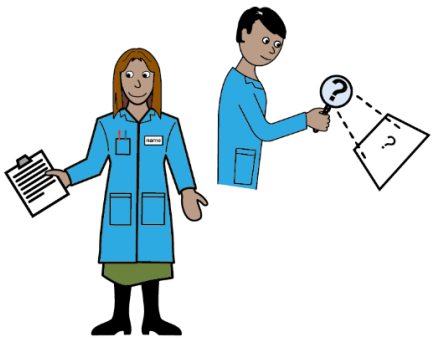
Participant's name:

Signature:

TESTING OBJECT ASSOCIATION

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

EASY READ PARTICIPANT INFORMATION SHEET

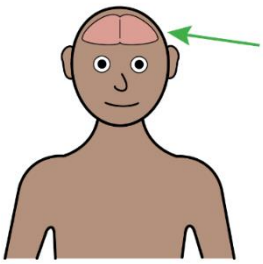


You are invited to take part in a study.

What is this study about?

This study is to find out if a new test for brain problems works better than the tests we have now.

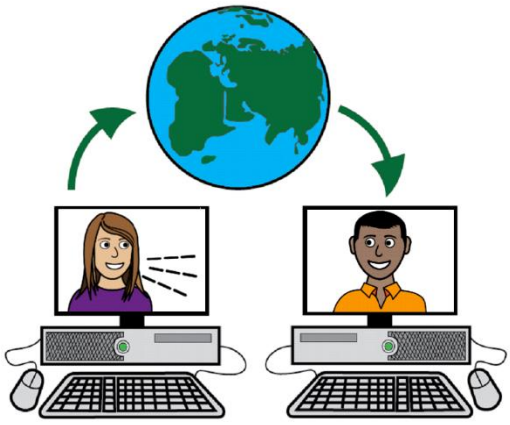
Who can take part in the study?



Adults who have the kind of brain problem that the tests are for.

Your Speech Language Therapist gave you this information because you have this kind of brain problem.

What will I do if I take part?



You and I would have a meeting by video.

It would take about 45 minutes. You can ask a support person to come to the meeting with you.

Your Speech Language Therapist might help you set up the meeting.

You can watch a video of what the meeting would look like at http://bit.ly/studymeeting

After the meeting, I would talk to your Speech Language Therapist about your brain problem. We would not talk about anything else.

What would happen in the meeting?

First, we would talk about the study and I would answer any questions. I might ask you some questions to check that I explained the study well.

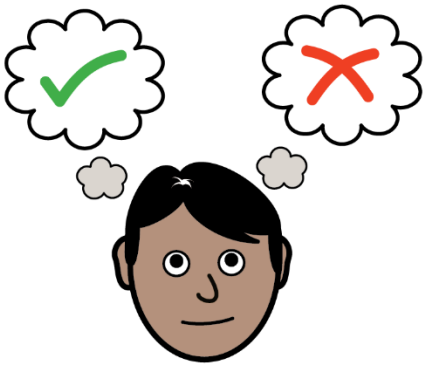
We would fill in a form together that says you want to take part. I would take a photo of you filling in the form.

Second, you would do 3 tests. In the tests, you would:

- look at pictures and choose which picture is correct
- look at pictures and choose which pictures go together
- think of words and say them as fast as you can

Last, we would talk about how the tests went.

Do I have to take part?



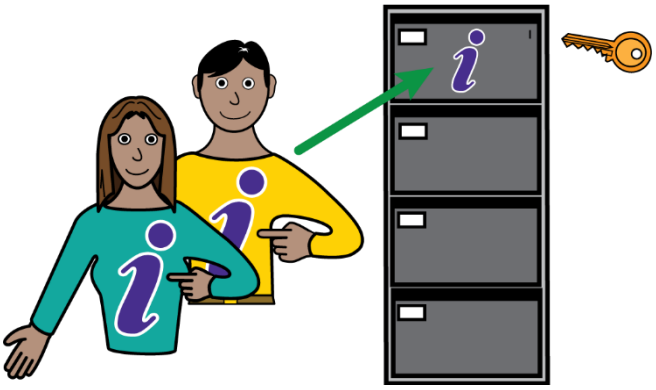
No. You do not have to do this study. It is your choice.

Taking part, or not, will not change the support you get from your Speech Language Therapist or any other services.

Can I change my mind?

Yes. You can change your mind at any time when we meet. I will stop the meeting. You do not have to tell me why you want to stop.

After we talk you can still change your mind. If you tell me you have changed your mind, I will not use any of your information.



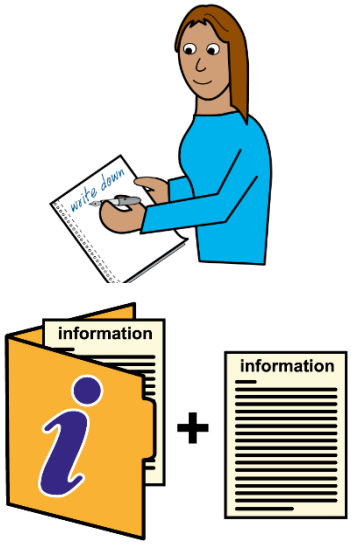
What happens to my information?

Your information will be kept private and safe.

Your test results will be protected with a password. They will not have your name on them.

You can ask to see the information we have about you. You can change the information if it is wrong.

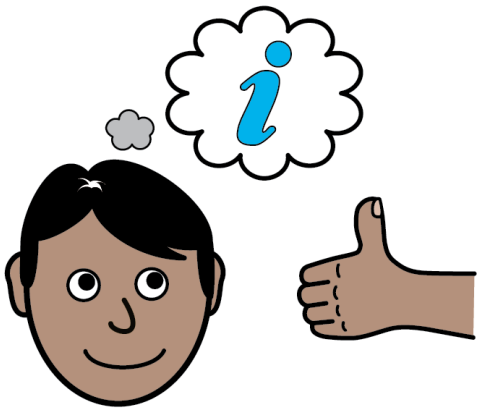
After the study ends, Massey University will keep your information safe for 10 years. Then they will destroy it.



What will you do with the results?

I will write about the results. I will send this to Massey University. I will also send it to a science journal so other people can read about it.

I can send you information about the results too if you like.

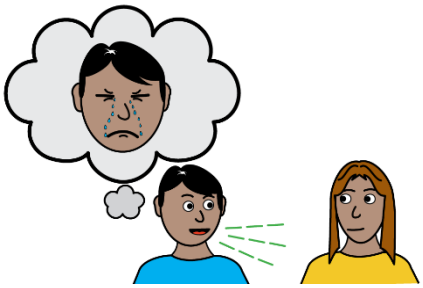


Good things about taking part?

You might think the tests are interesting.

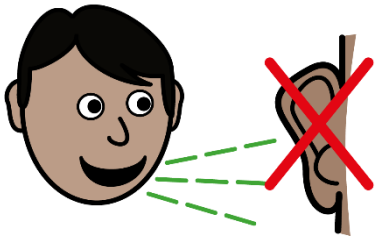
The study could help other people who are sick. You might feel happy that you took part.

You will get a $30 supermarket voucher.



Risks when taking part?

Some people may feel upset or worried if the tests are hard. Tell me if you feel upset or worried. We can talk about how you are feeling and I will try to help.



Someone could hear what we say in the meeting. I will be in a private place so no-one can hear me. You should choose a private place for the meeting too.

Because we will meet by video, someone else could listen to what we say online. I will use a computer program that makes it hard for someone to listen.

Who can I talk to about the study?

My name is Jo Chapman. I am doing this study for my degree at Massey University. You can talk to me if you have questions or concerns about the study:

Jo Chapman
Email: [email]
Phone: [phone number]

You can also talk to my supervisors, who are helping me. They are:

Professor Janet Leathem from Massey University
Email: [email]

Dr Stephen Hill from Massey University
Email: [email]

Dr Kwok-Keung Leung
Email: [email]

For Māori support, you can talk to a Kaimātai Hinengaro Matua: Māori:

Dr Simon Bennett from Massey University

Email: [email]

You may want to talk to someone who is not part of the study. You can talk to an **independent health and disability advocate** on:

Phone: 0800 555 050
Email: advocacy@advocacy.org.nz
Website: https://www.advocacy.org.nz/

The **New Zealand Health and Disability Ethics Committee** checked this study. They are a group of people who make sure studies are run well. You can contact them on:

Health and Disability Ethics Committees

Phone: 0800 4 ETHICS
Email: hdecs@moh.govt.nz

TESTING OBJECT ASSOCIATION

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

SPEECH LANGUAGE THERAPIST INFORMATION SHEET

Ethics committee reference: 20/NTB/93
Lead investigator: Jo Chapman

You are invited to assist with the validation of a new test of semantic association.

We are asking SLTs to identify clients who are eligible for the study, share information about the study with those clients, and facilitate clients' contact with the researchers. Your participation is voluntary and your clients are also free to choose whether or not to participate.

After reading this information, please contact us to register to participate. We appreciate you considering participation and also welcome you contacting us with any questions.

PURPOSE OF THIS STUDY

This study aims to validate a new test of semantic association.

Psychometric tests can be used during brain tumour surgery to help surgeons decide which parts of the brain can be safely removed. Tests of different cognitive functions are needed for different parts of the brain. The test of non-verbal semantic association currently used for this purpose is not ideal and Wellington Regional Hospital approached Massey University to develop a replacement. A better test could make tumour surgery safer and more effective.

We have already conducted a general population study to ensure our new test is easily completed by adults without any cognitive impairments.

This current study aims to determine whether the test is sensitive and specific to semantic impairment. People with specific semantic impairments are the ideal population for this validation. We are contacting SLTs as the group of professionals best positioned to identify people with these conditions. We are also separately recruiting patients with brain tumours, but many tumours don't cause any cognitive symptoms so their results may not be enough to determine the test's validity.

Jo Chapman is conducting this study as part of her Doctor of Clinical Psychology degree at Massey University. Her supervisors are Professor Janet Leathem and Dr Stephen Hill from Massey University and Dr Kwok-Keung Leung.

This study has been approved by the New Zealand Health and Disability Ethics Committee, reference 20/NTB/93.

ELIGIBILITY

Your adult clients may be eligible for this study if they have a **specific semantic impairment**. Possible causes include Primary Progressive Aphasia (semantic variant, also called Semantic Dementia) or a small localised stroke.

Clinical signs may include:

- Difficulty with naming and single-word comprehension
- Asking the meaning of previously familiar words e.g. "What is…?"
- Using generic imprecise words e.g. "animal" or "thing"
- Unimpaired grammar and motor speech

To ensure the test results are specific to semantic impairment, we are **not** looking for clients who also have:

- Any other significant language impairment
- A visual or perceptual impairment (not corrected by glasses)
- Any other significant neurological condition (e.g. traumatic brain injury, dementia)

WHAT PARTICIPATION WOULD INVOLVE

We will provide you with a Participant Information Sheet for your clients, in Easy Read format if appropriate. We ask that you discuss the information with them, give them an opportunity to ask questions, and provide them a copy of the document to take away and consider or discuss with others.

If your client wishes to participate in the study, we will meet with them via Zoom video conference. They can arrange this directly with us if they are comfortable doing so. Alternatively, you can support them in learning how to use Zoom and facilitate an initial Zoom meeting to 'introduce' us. Your client will need to have access to a computer with video/audio capability and a stable internet connection.

When we meet with your client, we will talk through the information sheet and answer any questions they have before asking for consent to participate. We will record consent using an online form which we will complete together, accompanied by an image of the participant. Participation will involve sharing some demographic and medical information with us and completing three tests: the new test and 2 existing tests for comparison. The tests will involve non-verbal semantic association and verbal fluency.

After meeting with your client, we will contact you (with their permission) to ask about the diagnosis causing their semantic impairment. We won't ask for other information about them or to see any of their medical records.

POSSIBLE BENEFITS AND RISKS

You will have contributed to a study that could improve outcomes for people with brain tumours. We will also send you a $30 supermarket voucher as a token of our appreciation.

For your clients, taking part in this study will not change their condition or treatment. They may find extra formal testing unpleasant and/or feel upset or worried if they find the tests difficult. They may also feel pleased to have taken part in a study that could help others. We will do our best to make participation a positive experience for your clients. We will also send them a $30 supermarket voucher to thank them for meeting with us.

YOUR AND YOUR CLIENTS' RIGHTS

You and your clients can choose whether or not to participate and can change your minds at any time during the study without giving a reason.

We will keep your clients' medical information and test results separate from their name and identifying details. However, their diagnosis and demographic details are likely to be unique in New Zealand so this information could potentially be used to identify them. The study data will be kept in a secure online database only accessible by the study team. Your clients can ask to see the information we have about them at any time.

All publications resulting from the study will report aggregated data only and will not include any information about individual participants.

WHAT WILL HAPPEN AFTER THE STUDY

If you wish, we will send you and/or your clients a summary of the results of the study at the end of 2021, after it is complete. The results of this study will form part of a Massey University Doctor of Clinical Psychology thesis and be published in relevant journals.

Massey University will keep the information from this study securely for 10 years, then destroy it.

WHO TO CONTACT ABOUT THE STUDY

To register to participate or if you have questions, concerns or complaints about the study at any stage, please contact the study team:

Jo Chapman	[email]
	[phone number]
Professor Janet Leathem	[email]
Dr Stephen Hill	[email]
Dr Kwok-Keung Leung	[email]

You can also contact the health and disability ethics committee that approved this study on:

Phone:	0800 4 ETHICS
Email:	hdecs@moh.govt.nz

TESTING OBJECT ASSOCIATION

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

RECRUITMENT QUICK REFERENCE

Ethics committee reference: 20/NTB/93
Lead investigator: Jo Chapman – [phone number]

GLIOMA	SEMANTIC IMPAIRMENT
Eligible patients: • Current *OR* previous glioma • Over 18 Ideal patients: • Have had a DTI scan • You can assess whether their IFOF was invaded, displaced, or spared by their tumour	**Eligible** patients: • Specific semantic impairment (e.g. svPPA, small stroke) • Over 18 Clinical signs may include: • Difficulty with naming and single-word comprehension • Asking the meaning of previously familiar words e.g. "What is…?" • Using generic imprecise words e.g. "animal" or "thing" • Unimpaired grammar and motor speech
Ineligible if they have: • Any other significant language impairment • A visual or perceptual impairment (not corrected by glasses) • Any other significant neurological condition (e.g. traumatic brain injury, dementia)	
Tell patient they are eligible for a study about a new test. **Tell** them participation is optional and will not affect their care. **Give** patient the appropriate information sheet: • Glioma – current • Glioma – previous • Semantic • Semantic – EasyRead (also give Semantic information sheet for family) **Ask** patient if they would prefer to contact the researcher themselves *or* for the researcher to contact them. **Tick** the appropriate box on the front of the information sheet. **If** the patient prefers the researcher to contact them, **email** the patient's name and contact details to: Joanne.Chapman.2@uni.massey.ac.nz **Do not** email medical information at this stage.	

Appendix G: Clinical Validation Meetings

This appendix contains the following documents, related to conducting meetings with participants in the clinical validation study (described in chapter 6):

1. Study meeting guide, with test administration instructions and distress management plan
2. Screenshots of Qualtrics surveys used

Study meeting procedure guide

Opening and consent

Small talk, check sound/video quality, introductions.

Display appropriate information sheet.

Say: **You've already seen this, but I have to talk through it again with you. Feel free to ask me questions at any point as we go through.**

Regularly say: **Does that make sense? Is that ok with you? Do you have any questions about that?**

Key points to cover:

1. Participation is your choice
2. We are trying to find out if a new test works
3. Agenda for this meeting
    a. Fill in consent form
    b. Take a picture
    c. Demographic questions
    d. 3 tests
    e. Discussion
4. Contact with doctor, information requested
5. Benefits and risks – discuss each paragraph
6. Sending results
7. Storing information
8. Contacts

Say: **I want to check that I explained that properly. Can you tell me what you understood?**

If needed, use specific questions to check key points e.g. **What will we do today? What will I ask your doctor/therapist about?**

If needed, use binary questions (including both answer options) e.g. **Do you have to do the tests or can you say no? After today, will your therapy change or stay the same?**

Say: **Do you want to go ahead and take part in this study?**

Display consent form and give participant control.

Say: **Do you have any trouble reading? Some people prefer I read the form out loud.**

Discuss / Assist with form as needed.

Take screenshot after form signed, save to folder, name with date and time.

Fill in researcher section of form and submit.

Display demographics page.

Say: **First I need some basic information about you. This is so we can check whether the tests work with different kinds of people.**

Ok. Now let's start the tests.

Pyramids and Palm Trees Test

Show first practice item.

First test	Last test
This is the first test.	**This is the last test. It is a bit like the test we did first.**

Say: **Here are three pictures. You have to decide which one of these two at the bottom goes with the one at the top. Is it this one or this one?**

Correct	Incorrect
That's right, they go together because a waistcoat and a bow tie are both worn by men.	**No. The bow tie goes with the waistcoat because they are both worn by men.**

Show next practice item.

Say: **Now try this one. Which of these two pictures goes with the one at the top?**

Correct	Incorrect
That's right, they go together because you pour from a bottle into a glass.	**No, it's this one. You normally use a glass with a bottle.**

Show next practice item.

Say: **Now try this one. Which of these two pictures goes with the one at the top?**

Correct	Incorrect
That's right, they go together because both a clown and a lion can be found at a circus.	**No, it's this one. They go together because you find both a clown and lion at a circus, but not a giraffe.**

Show first test item.

Say: **Now you do the rest. After you choose an answer, the next question will appear automatically. Go ahead.**

If examinee is having difficulty or is frustrated, say: **It's like a game of guessing what connection the test maker had in mind.**

If examinee says the two bottom images go together, say: **You have to choose which of these two goes with the one at the top.**

If examinee is unsure, strongly encourage them to guess.

New test

Show practice item.

First test	Last test
This is the first test.	**This is the last test. It is a bit like the test we did first.**

Say: **Here are three pictures. You have to decide which picture is correct. Which one of these do you think is correct?**

Say: [**That's right.** / **No, it's this one.**] **The sun and a cloud can appear together in the sky but neither pencils nor fridges are likely to be in a cloud.**

All the test questions are similar. For each question, choose which picture you think is correct. If you're not sure, take a guess. When you select your answer, the next question will appear automatically.

Any questions?

Show first test item.

Say: **Here is the first question. Go ahead.**

If examinee says multiple pictures could be correct, say: **Maybe, but which one is the *most likely*?**

If examinee asks what a picture is, say: **Just answer based on what it looks like to you.**

Verbal fluency

Have timer, pencil, and paper ready.

Say: **Now the next one is a bit different.**

I will say a letter of the alphabet. Then I want you to give me as many words that begin with that letter as quickly as you can. You'll have one minute.

For example, if I say 'G' you might give me 'golf, grey, gamble, go…'

I do not want you to use words that start with a capital letter, such as 'George' or 'Greytown'.

Also, do not use the same word with different endings such as 'go', 'going', 'goes'.

Here's a reminder of those rules. Any questions?

So, if I said 'H' what could you say? [Feedback as appropriate] [Horse, huge, hang, honest, hand…]

Ok. Begin when I say the letter. The first letter is 'F'. Go ahead.

Record words as the examinee says them.

If a silence of 15 seconds, repeat the basic instructions and the letter.

If examinee stops, encourage them to try to think of more words.

At the end of the minute, say **Fine** or **Good**.

Say: **Now we'll do it again with a different letter. Ready? The next letter is 'A'. Go ahead.**

Repeat for letter **S**.

Backup letters if something goes wrong during administration: **P**, **L**

Say: **Just one more. Ready? This time, tell me the names of as many *animals* as you can. It doesn't matter what letter they start with. Go ahead.**

Backup categories if something goes wrong during administration: **vegetables**, **body parts**

Managing distress

If a participant expresses distress during the debrief after testing:

- Listen supportively.
- Validate and normalise difficulty with the tests e.g. **"Yes, these tests are tricky! This is because we did just the ones that pick on the thing you have trouble with." "Lots of people find that one hard."**
- Remind participants that their other cognitive abilities are unaffected e.g. **"You're able to have this conversation with me, [specific capability examples e.g. work, hobbies, family]. If I gave you other tests where you drew shapes or did maths or found patterns you would probably do really well."**
- Reframe difficulty as helping the study e.g. **"You having trouble with these is actually really good, it tells me the tests are working! This way I can get the best information about what happens when people have trouble with them."**
- Reframe difficulty as a problem with the test e.g. **"This is a very new test, we might decide those questions are too confusing and stop using them."**

If a participant appears distressed during testing:

- Use encouraging statements, validate/normalise difficulty, and praise effort e.g**. "you're doing well", "that was a tricky one", "let's try another one", "you're working so hard on these", "thank you for all the effort you're putting into this", "just do your best"**.

If a participant's distress during testing escalates and/or does not reduce:

- Pause testing.
- Check in with participant e.g. **"Are you ok to keep going or would you like a break?", "What came up for you just then?"**
- Follow debrief process as above.
- Offer choice whether to resume test, or move on to next test, or end the session.
- If participant wishes to end the session, accept and reassure e.g. **"No problem. Thank you for having a go. I really appreciate you taking the time to meet with me."**

If a participant appears extremely distressed:

- Stop testing.
- Acknowledge and accept distress, offer time out e.g. **"It's ok. Let's take a break for a bit."**
- Allow silence, model deep breathing.
- Use simple empathic statements e.g. **"That was really tough, wasn't it?"**
- Coach participant through paced breathing and/or suggest getting a drink.
- Ask participant if they would like to talk about what happened.
- Follow process as above.

Closing and feedback

Say: **That's all the tests done. How did you find that?**

Key points to cover:

1. Any specific comments on the tests
2. Are they worried about not doing well or finding it hard?
3. Any questions?
4. Which test was which
5. Do they want to know how they did? – can email later
6. Thank you

TESTING OBJECT ASSOCIATION

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

CONSENT FORM

Please tick to indicate you consent to the following:

- I have read, or have had read to me, and I understand the Participant Information Sheet. I have a copy of the information sheet and this form. ☐
- I have been given enough time to consider whether to take part in this study. I have had the opportunity to use a support person to help me ask questions and understand the study. ☐
- I am satisfied with the answers I have been given about the study. I know who to contact if I have any further questions. ☐
- I understand that taking part in this study is my choice and that I may withdraw from the study at any time without affecting my medical care. ☐
- I consent to the study team collecting and processing my information, including specific information from my doctor or therapist about my health. ☐
- I understand that my information is confidential and that no material which could identify me will be used in any reports on this study. ☐
- I understand the video conferencing software is secure, but privacy online cannot be guaranteed. ☐
- I consent to the researcher recording an image of me and this form as evidence of my consent. ☐

I wish to receive a summary of the results from the study.

Yes ○

No ○

If yes, please send the summary of results to this email address:

Please send my $30 supermarket voucher to this address:

I consent to take part in this study.

Participant's name:

Signature:

SIGN HERE

clear

Date:

27 Aug 2021

I have given a verbal explanation of the research project to the participant and have answered the participant's questions about it. I believe that the participant understands the study and has given informed consent to participate.

Researcher's name:

Signature:

SIGN HERE

clear

Done

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

First, please tell us some information about yourself.
This helps us to check whether the new test will be able to used with a wide variety of people.

What is your age?

- ○ 18 - 29
- ○ 30 - 39
- ○ 40 - 49
- ○ 50 - 59
- ○ 60 - 69
- ○ 70 - 79
- ○ 80 or older

What gender do you identify as?

- ○ Male
- ○ Female
- ○ Nonbinary / Genderfluid / Other

What race or ethnicity do you identify with?
Select all that apply.

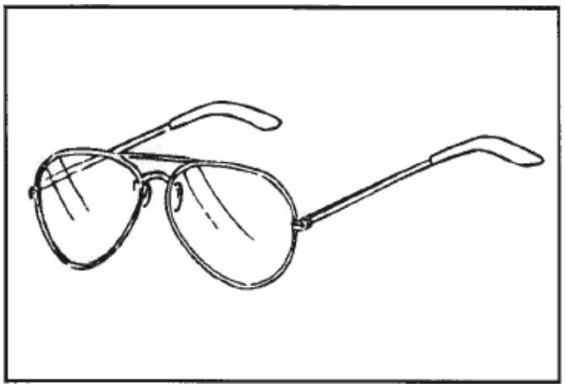
- ☐ White / European / Caucasian
- ☐ NZ Māori or Pacific Islander
- ☐ Latino / Hispanic
- ☐ East Asian or Asian Indian
- ☐ South Asian or Indian
- ☐ Middle Eastern or North African
- ☐ African
- ☐ African American
- ☐ American Indian or other North American Native
- ☐ Other

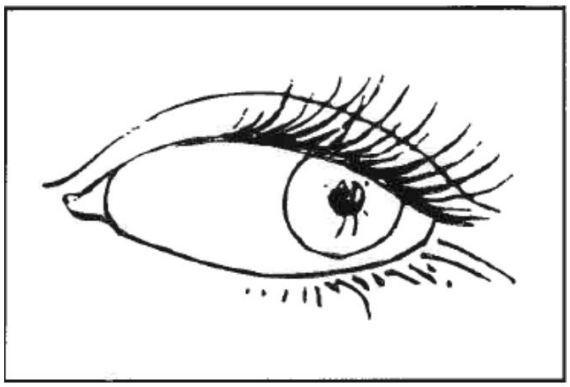
What is the highest educational qualification you have received?

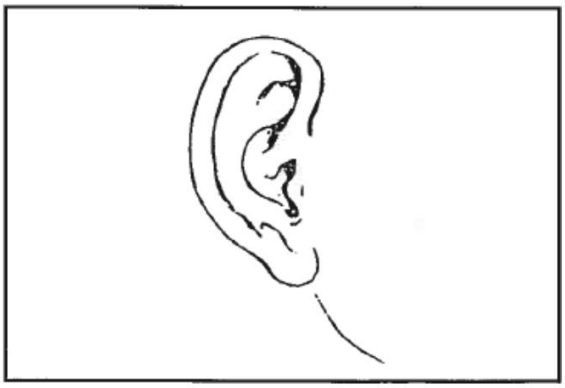
- ○ no formal qualification
- ○ a secondary school qualification (high school)
- ○ completed secondary school (college/university entrance or equivalent)
- ○ trade/technical/vocational training
- ○ undergraduate degree (bachelors or equivalent)
- ○ postgraduate degree (masters, doctorate or equivalent)

→

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND







○ ○

←

→

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

Well done. You have completed the first test.

←

→

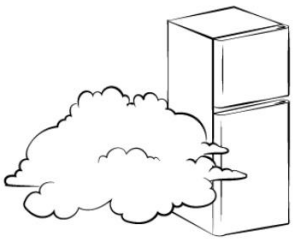

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND
Words that start with the letter.
Not words that start with a capital letter.
Not the same word with different endings.
←
→


MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND
Names of **animals**.
←
→


MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND
Well done. You have completed the second test.
←
→

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

Select the correct image.



→


MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND
Well done. You have completed all the tests.
→

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

Notes/observations from test administration: New Test

Notes/observations from test administration: PPTT

Notes/observations from test administration: Verbal Fluency

Participant comments: New Test

Participant comments: PPTT

Participant comments: Verbal Fluency

Participant comments: Other

←

→

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

Verbal fluency: responses

F

A

S

Animals

Verbal fluency: scores

F

A

S

Animals

←

→

Appendix H: Conference Posters

This appendix contains the two posters presented at conferences during this project:

- 37th International Australasian Winter Conference on Brain Research, 2019
- 2021 Neurological Association of New Zealand Meeting

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

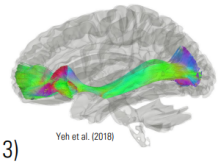
A NEW TEST DESIGN
FOR IFOF MAPPING DURING AWAKE CRANIOTOMY

J.C. CHAPMAN[1], J.M. LEATHEM[1], and K.-K. LEUNG[2]
[1]School of Psychology, Massey University, Wellington, NZ [2]Department of Neurology, Wellington Regional Hospital, NZ

INFERIOR FRONTO-OCCIPITAL FASCICULUS
(IFOF)

Part of ventral language pathway, also involved in:
- Semantic memory
- Emotion
- Prosody
- Social cognition
- Spatial neglect

(Duffau, Herbet, & Mortiz-Gasser, 2013)

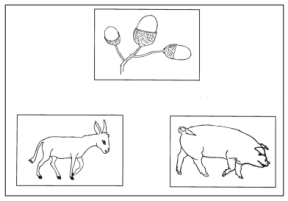

Yeh et al. (2018)

AWAKE CRANIOTOMY
- Increasingly used in brain tumour treatment
- Individualised brain mapping through the administration of tests
  - Includes both cortical regions and sub-cortical tracts, such as the IFOF
  - Procedures for specific regions or tracts are not standardised and vary widely
- Enables increased tumour resection and reduced post-operative deficits

(Talacchi et al., 2013)

CURRENT ASSESSMENT
Pyramids And Palm Trees Test

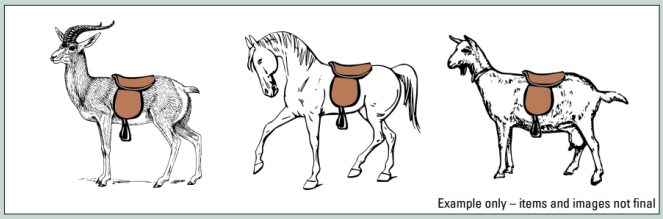
Which of the lower pictures goes with the top one?



Problems:
- Out of date objects
- Cultural bias
- Education effects
- Low image quality
- Low reliability and validity
- No theoretical background or definition of construct

(Breining et al., 2015; Klein & Buchanan, 2009; Olivarez & Boroda, 2007)

NEW TEST
Composite Images



Point to the correct one

Format advantages:
- Simple instructions and single point response
- Language independent
- Requires differentiation of specific object features

Other Semantic Tests
- Too narrow in focus for IFOF mapping
- Too long for surgical use
- Require category knowledge only
- Language dependent

Item selection is based on:
- Equal items from each category
- Low age of acquisition
- High familiarity
- Strength of association (using latent semantic analysis)

NEXT STEP: Validation
- General population - check for culture, age, or education effects
- Clinical groups – patients with glioma (with and without IFOF involvement) or Primary Progressive Aphasia (semantic variant)

DO YOU HAVE PATIENTS WITH GLIOMA OR SV-PPA?
Ask us how they can be involved: joanne.chapman.2@uni.massey.ac.nz

REFERENCES

- Breining, B. L., Lala, T., Cuitiño, M. M., Manes, F., Peristeri, E., Tsapkini, K., … Hillis, A. E. (2015). A brief assessment of object semantics in primary progressive aphasia. *Aphasiology, 29*(4), 488–505.
- Duffau, H., Herbet, G., & Moritz-Gasser, S. (2013). Toward a pluri-component, multimodal, and dynamic organization of the ventral semantic stream in humans: Lessons from stimulation mapping in awake patients. *Frontiers in Systems Neuroscience, 7.*
- Klein, L. A., & Buchanan, J. A. (2009). Psychometric properties of the Pyramids and Palm Trees Test. *Journal of Clinical and Experimental Neuropsychology, 31*(7), 803–808.
- Olivarez, A., & Boroda, A. D. (2007). Review of The Pyramids and Palm Trees Test. In K. F. Geisinger, R. A. Spies, J. F. Carlson, & B. S. Plake (Eds.), *The seventeenth mental measurements yearbook.*
- Talacchi, A., Santini, B., Casartelli, M., Monti, A., Capasso, R., & Miceli, G. (2013). Awake surgery between art and science. Part II: Language and cognitive mapping. *Functional Neurology, 28*(3), 223–239.
- Yeh, F. C., Panesar, S., Fernandes, D., Meola, A., Yoshino, M., Fernandez-Miranda, J. C., … & Verstynen, T. (2018). Population-averaged atlas of the macroscale human structural connectome and its network topology. *NeuroImage, 178*, 57-68.

POSTER PRESENTED AT: 37th International Australasian Winter Conference on Brain Research, Queenstown, New Zealand, 31 August – 4 September 2019

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

A NEW TEST
FOR IFOF MAPPING DURING AWAKE CRANIOTOMY

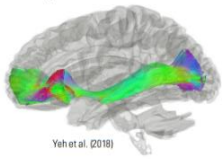
J.C. CHAPMAN and J.M. LEATHEM
School of Psychology, Massey University, Wellington, NZ

INFERIOR FRONTO-OCCIPITAL FASCICULUS (IFOF)

Part of ventral language pathway, also involved in:

- Semantic memory
- Emotion
- Prosody
- Social cognition
- Spatial neglect

(Duffau, Herbet, & Mortiz-Gasser, 2013)



AWAKE CRANIOTOMY

- Gold standard for glioma treatment
- Individualised brain mapping through the administration of tests
  - Both cortical regions and sub-cortical tracts
  - Procedures are not standardised
- Enables increased tumour resection and reduced post-operative deficits

(Gogos et al., 2020; Talacchi et al., 2013)

CURRENT IFOF MAPPING TOOLS

Pyramids and Palm Trees Test (PPTT)

- Out of date objects
- Cultural and educational bias
- Low reliability, validity, and image quality

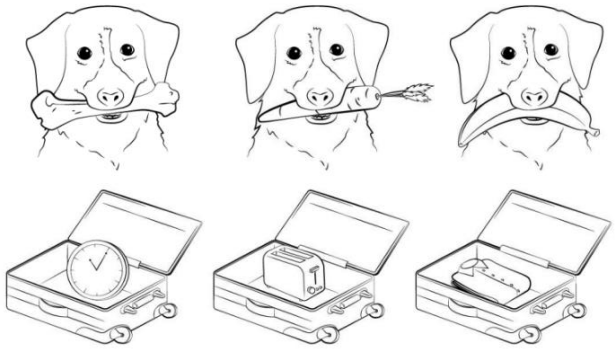
(Klein & Buchanan, 2009; Olivarez & Boroda, 2007)

Other semantic tests

- Too narrow in focus for IFOF mapping
- Too long for surgical use
- Assess category knowledge only
- Language dependent

NEW TEST – DESIGN
HIGHER QUALITY, MORE CONSISTENT

Which image is correct?



New format:

- Simple instructions and single point response
- Language independent
- Requires differentiation of specific object features

Item selection based on:

- Low age of acquisition and high familiarity
- Strength of association (using latent semantic analysis)

Quality Measure	PPTT mean (SD)	New Test mean (SD)	
Word familiarity rating	5.9 (0.71)	6.4 (0.25)	*p<.001*
Word age of acquisition rating	2.6 (0.81)	2.2 (0.36)	*p<.001*
Associative strength of item targets	.33 (.15)	.46 (.09)	*p<.001*
Associative strength of item distractors	.22 (.15)	.21 (.08)	*p=.52*

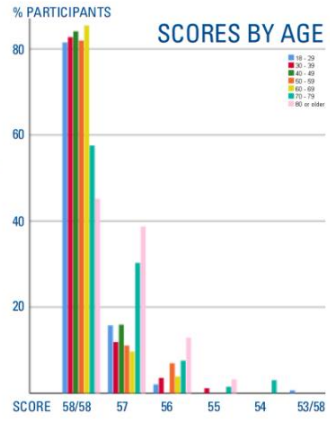
NEW TEST - VALIDATION
LESS BIAS

Trials with 707 healthy adults

Item performance evaluated:

- Incorrect answers (ideally none)
- Time spent per item (ideally short)
- Participant comments (ideally none)

Removed or updated items not performing as designed



Group score differences:

No differences between:

- Genders
- Levels of education
- Ages 18-69
- European, Māori/Pasifika, or Asian ethnicities

Slightly lower scores:

- Ages 70+, 'Other' ethnicities

CLINICAL VALIDATION
SENSITIVE AND SPECIFIC?

Trials in progress with relevant patient groups:

- Glioma
- Semantic impairment

DO YOU HAVE PATIENTS WITH SV-PPA OR SEMANTIC IMPAIRMENT?

Ask us how they can be involved: joanne.chapman.2@uni.massey.ac.nz

REFERENCES

Duffau, H., Herbet, G., & Moritz-Gasser, S. (2013). Toward a pluri-component, multimodal, and dynamic organization of the ventral semantic stream in humans: Lessons from stimulation mapping in awake patients. *Frontiers in Systems Neuroscience, 7.*
Gogos, A. J., Young, J. S., Morshed, R. A., Hervey-Jumper, S. L., & Berger, M. S. (2020). Awake glioma surgery: Technical evolution and nuances. *Journal of Neuro-Oncology, 147*(3), 515–524.
Klein, L. A., & Buchanan, J. A. (2009). Psychometric properties of the Pyramids and Palm Trees Test. *Journal of Clinical and Experimental Neuropsychology, 31*(7), 803–808.
Olivarez, A., & Boroda, A. D. (2007). Review of The Pyramids and Palm Trees Test. In K. F. Geisinger, R. A. Spies, J. F. Carlson, & B. S. Plake (Eds.), *The seventeenth mental measurements yearbook.*
Talacchi, A., Santini, B., Casartelli, M., Monti, A., Capasso, R., & Miceli, G. (2013). Awake surgery between art and science. Part II: Language and cognitive mapping. *Functional Neurology, 28*(3), 223–239.
Yeh, F. C., Panesar, S., Fernandes, D., Meola, A., Yoshino, M., Fernandez-Miranda, J. C., ... & Verstynen, T. (2018). Population-averaged atlas of the macroscale human structural connectome and its network topology. *NeuroImage, 178*, 57-68.

POSTER PRESENTED AT 2021 Neurological Association of New Zealand (NANZ) meeting, Wellington, New Zealand, 4-5 November 2021

Appendix I: Research Case Study

Research Case Study: Psychologists in Multi-Disciplinary Teams

This research case study represents the work of Jo Chapman during the period of an internship in partial fulfilment of the degree of Doctor of Clinical Psychology at Massey University.

In accordance with the Code of Ethics for Psychologists Working in Aotearoa/New Zealand, client confidentiality is maintained by using pseudonyms and adapting identifying information.

Candidate:	Jo Chapman
Placement:	Community Mental Health and Addictions (South)
	Te Whatu Ora Health New Zealand Te Matau a Māui Hawkes Bay
Supervisors:	Dr Inez Awatere-Walker, Consultant Clinical Psychologist
	Mr Ian Clayton, Senior Clinical Psychologist
Research Supervisor:	Professor Janet Leathem, Consultant Neuropsychologist

Word count: 3108

Abstract

This research case study demonstrates how my doctoral research influenced my practice as an intern psychologist through an exploration of multidisciplinary team working in both surgical and mental health settings. My research aimed to develop a psychological test for use during awake brain surgery, in collaboration with Neurology staff. I observed the medical team conducting two of these surgeries and my clinical placements and internship gave me opportunities to observe multidisciplinary teams in various mental health services. I noticed that members of well-functioning teams respected each other's expertise and consulted frequently, valuing both the process and each profession's contributions. A key role of Psychologists was in formulation, providing alternative hypotheses and assisting the team to 'make sense' of clients' presentations. Psychologists also acted as facilitators of reflective practice and support for the wellbeing of the team. Both clients and staff benefit from good multi-disciplinary team processes. In my future practice, I want to continue working as an integral member of a team, modelling and advocating for collaborative practice.

Psychologists in Multi-Disciplinary Teams

This case study presents a discussion of the roles of Psychologists in multi-disciplinary teams (MDTs), informed by my experiences during doctoral research, clinical placements, and internship. It will outline the background to my research project, followed by my observations of Clinical Psychologists working in surgical and mental health teams, and then briefly connect these observations to published literature. The final reflection section highlights my learnings and aims for future practice.

Doctoral Research Experiences

My doctoral research project, "Mapping the IFOF with a New Test of Semantic Association", aimed to develop and validate a new psychometric test to replace one currently used during awake surgery for brain tumours.

Background to the Project

Like most cancers, the aim of surgery for tumours arising from support cells in the brain (gliomas) is to remove as much of the tumour as possible. Total removal has been shown to improve survival times over subtotal or partial removal (Nitta & Sato, 1995). However, this aim must be moderated by avoiding any permanent severe cognitive deficits due to the loss of functioning brain tissue (De Witt Hamer et al., 2012). Higher grade tumours often have poorly defined borders and infiltrate healthy tissue making it very difficult for surgeons to determine which tissue needs to, or can, be removed and which does not.

To mitigate this problem, awake craniotomy is used for tumours located near critical brain regions. This is surgery to remove brain tissue performed with the patient awake and using only local anaesthetic, which is generally well tolerated by patients (Gernsback et al., 2018). During an awake craniotomy the patient is asked to perform a task while the surgeon uses a probe to 'map' which brain areas are involved in the task based on whether the patient's performance is impaired during stimulation of each area. Tumour removal begins after mapping of the brain's surface is complete and further mapping continues during removal as new areas are exposed (Talacchi et al., 2013). Thus, the locations of functional areas can be determined in each individual patient and tumour tissue can be removed precisely up to their borders. A meta-analysis by De Witt Hamer et al. (2012) found that awake craniotomy for gliomas produced less severe long term deficits and more complete tumour removal than anaesthetised surgery. They recommended awake craniotomy be implemented as the universal standard of care for glioma.

However, the safety and effectiveness of awake craniotomy depends on the accuracy of brain mapping. Tests used for mapping vary between hospitals and are generally selected based on the location of the tumour (Talacchi et al., 2013). Long-distance white matter tracts in the brain generally define the borders of tumour removal as they often perform complex functions which need to be preserved. These tracts are therefore particularly important to map accurately.

The Inferior Fronto-Occipital Fasciculus (IFOF) is one such tract with wide ranging connections throughout the brain (Bajada et al., 2015). It is involved in a variety of functions, including language, emotional functioning, visual perception, and semantic memory (Conner et

al., 2018). Semantic memory refers to the memory for general knowledge such as the names, properties, functions, and meanings of objects (Chen et al., 2017). It appears to be the most viable test target for IFOF mapping during surgery as it can be quickly tested with relatively simple tasks, unlike emotional or linguistic functioning (Lemée et al., 2018).

Naming tasks, in which a patient is shown an image of an object and asked to say its name aloud, are widely used and well regarded in mapping language-related brain areas, including the IFOF (Talacchi et al., 2013). However, the spoken response can be difficult for patients with a language impairment. In contrast, the Pyramids and Palm Trees Test (PPTT) is a non-verbal semantic test which is used to map the IFOF in awake craniotomy at Wellington Regional Hospital and elsewhere (Chang et al., 2018). Patients are asked to determine which images 'go together' and thus need to access and integrate more semantic information than during a naming task. This increased complexity is appropriate when mapping a tract like the IFOF, which integrates multiple brain areas. However, reviews of the test have noted it lacks a clear theoretical background or construct definition, the image quality varies widely, and many objects depicted are out of date (Olivarez & Boroda, 2007). International studies consistently find that people with fewer years of education receive lower PPTT scores, as do individuals from non-British cultural backgrounds (Callahan et al., 2010; Gamboz et al., 2009; Gudayol-Ferré et al., 2008; Klein & Buchanan, 2009). These findings raise concern about the PPTT's validity in these populations.

Wellington Regional Hospital approached Massey University to discuss developing a replacement for the PPTT. A test designed specifically for use in awake craniotomy, able to be used with patients from a wide variety of backgrounds, and both specific and sensitive to IFOF

impairment, could improve the guidance available to neurosurgeons and therefore outcomes for glioma patients.

The need for this research project, and awake craniotomy surgery itself, clearly illustrates the need for both surgical and psychological expertise in achieving optimal outcomes for brain tumour patients: knowledge of both anatomy and function; pathology and cognition; skills with surgical tools and psychometrics; supporting the patient's physical and psychological comfort. Further, academic researchers are needed to develop appropriate psychometric tests.

Multi-Disciplinary Teams in Surgery

I was privileged to observe two awake craniotomy surgeries to help me understand the context and requirements of the new test, and, during these, the crucial roles of speech language therapists, anaesthetists, nurses, and assistants also became evident. I noticed the unique contribution of each member of the team was respected and discussions or checks were frequent. At the same time, clinicians were willing to step outside their usual roles in response to the patient's immediate needs, doing whatever was required to support the overall goal. Psychologists' contributions to patients' medical treatment also occurred before and after surgery through pre-assessment of cognitive functioning, selection of relevant tests based on brain regions impacted by the tumour, preparing the patient and their family for the surgery process, assessment of post-surgical cognitive changes, and supporting the patient and their family to manage any ongoing impairments. My hospital supervisor encouraged this reflection by highlighting how knowledge was shared between team members, especially the consultations between Speech Language Therapists and Neuropsychologists in the selection of tests and interpretation of the patient's responses. We discussed the different areas of

expertise of each profession and how each small interaction during the process supported the development of a collaborative understanding and plan for the patient's care.

Multi-Disciplinary Research

Developing the new test itself required the expertise of many different professions. At various stages of the project I depended on or benefitted from the work of:

- Psychologists – psychometric theories, semantic memory theories, existing semantic tests, research into conditions affecting semantic memory, supervision
- Surgeons – anatomical understandings of the IFOF, practicalities of surgery, characteristics and behaviour of tumours
- Neurologists – functional understandings of the IFOF, effects of stimulation
- Speech Language Therapists – semantic memory theories, comprehension and language assessments, research into conditions affecting semantic memory
- Educators – principles of test design
- Linguists – theories of semantics and semantic association, psycholinguistic data sets
- Mathematicians – mathematical models of semantic association, statistical methods for psychometric design and evaluation
- Software Developers – computerised implementation of mathematical models, calculation and search tools, survey software
- Artists – collaborating with me to produce clear and appropriate image stimuli

I also needed to be somewhat multi-disciplinary myself as my professional background as a Software Developer was integral to the process of producing the new test. I built databases and wrote quite a bit of computer code during the project and find it difficult to

imagine how I would otherwise have done it. My supervisor told me the best research happens when people bring many different skill-sets together to solve a problem.

Placement and Internship Experiences

I had the opportunity to observe and participate in multi-disciplinary team meetings (MDTs) across various mental health settings during my external placements and internship. Each service had rules defining when clients should be presented at an MDT, generally including initial assessment/intake, discharge, regular review (e.g. 3 monthly), and whenever there was acute risk or concern. While some of these discussions were essentially simple sign-offs on decisions, others aimed to select the most appropriate clinicians to work with a given client or collectively generate ideas in difficult situations. One service also regularly convened 'complex MDTs', one-off meetings of a selected group of clinicians to formulate a plan for a specific client. Team roles and procedures for agenda setting or recording minutes varied widely and some of these MDTs were clearly valued and appreciated by team members while others appeared less effective.

These teams included social workers, addiction counsellors, and psychiatrists, rather than surgeons, anaesthetists, and researchers, but arose from the same need to integrate the knowledge of different professions to achieve the best outcome. In well-functioning mental health MDTs, I noticed many of the characteristics I had seen in surgical teams: respect for the contributions of each profession, frequent consultation, and willingness to step outside assigned roles when needed. Clients benefited from the knowledge of professions beyond their assigned key workers' as team members contributed advice, suggestions, or interpretations.

Questions from others encouraged reflective practice or identified areas for further assessment. The personal knowledge of team members was also often of value (e.g., identifying unsafe streets) as were personal relationships with clinicians in other services. The role of humour and social support became evident as MDTs provided a safe space to express stress, anxiety, disappointment, or frustration.

Where MDTs seemed less effective, there sometimes appeared to be 'self-fulfilling prophecies' in operation. Team members felt MDTs were merely a 'tick-box exercise' required by management, so they presented cases quickly and did not discuss other presentations, thus not receiving any benefit from the process and reinforcing their original belief. New team members entering such an MDT would soon develop the same belief based on their experience. Service staffing demands could encourage the development of this situation by rewarding 'efficiency', penalising staff when 'not enough' cases were covered during a given meeting and, in effect, prioritising quantity over quality. In other cases, power imbalances between team members appeared to disrupt or shut down discussions. This was particularly noticeable when Psychiatrists were operating as leaders, rather than members, of the team and others felt unable to challenge them.

Restructures or other changes in services gave me the opportunity to participate in some MDT discussions about the MDTs themselves. Comments from other staff demonstrated how highly valued good MDTs are by all involved, and both reinforced my own observations and deepened my understanding. Nurse preceptors encouraged student nurses to notice both the knowledge shared between professions and the benefits of peer social support. While management sometimes suggested staff move between MDT groups in a service for fresh

perspective, team members found the relationships and trust developed in a stable group to be more important. Where meeting dynamics varied between MDTs, staff were particularly reluctant to move from an MDT where discussions were expected to one that prioritised 'efficiency'.

The relationships developed within MDT meetings carry on outside that context and can continue to improve practice in other ways. Clinicians learn about each other's areas of expertise or interest and thus know who might be best to ask when specific issues arise. Pre-existing relationships made it much easier for me to approach someone with a question and I observed other staff approached me in the same way.

The Role of Psychologists

The shortage of Psychologists often means we are only able to work directly with a small proportion of clients with the most severe presentations. However, MDTs can provide an opportunity for more clients to benefit from psychological input. I have noticed Psychologists offering formulations to explain clients' presentations, advice on principles of behaviour change, suggestions of approaches to try in difficult interactions, and psychoeducation on the effects of trauma or brain injury. In many teams, this contribution was highly valued and explicitly sought by keyworkers, who seemed to appreciate Psychologists' ability to 'make sense' of client presentations. In some of my early placement experiences I was surprised by the apparent simplicity of discussions about complex clients, reminding me that clinical students tend to forget how much we know and reinforcing the importance of sharing expertise. (Further experience clarified that this 'knowledge gap' goes both ways and we have just as much to learn from other disciplines.) My supervisors have again encouraged this

reflection, emphasising that communicating with and advising colleagues is a key psychological skill.

I have also observed Psychologists taking the lead when members of the team were distressed, using clinical skills to explore and validate their experience, reframe difficulty, and support problem solving. Mental health work can be highly stressful, especially in the current environment of generally under-staffed services, and this peer-support benefits not only the clinicians involved but also, through them, their clients.

Connections to Literature

Research suggests well-functioning MDTs improve both client outcomes and staff wellbeing (Onyett, 2007). Poorly functioning teams, however, appear to produce worse outcomes than no multi-disciplinary input at all (Onyett, 2007). Improved clinician wellbeing is associated with improved client outcomes and staff relationships can play an important role in preventing burnout or compassion fatigue (Nutt & Keville, 2016). This supports my observations of social support in MDTs and other clinician's comments about the value of trusted relationships in stable teams.

In a report prepared for the British Psychological Society, (Onyett, 2007) explored the roles of Psychologists in multi-disciplinary mental health teams, highlighting the benefits of collaboration as well as the importance of maintaining a unique professional identity. A key identified role was as a counter to the prevailing 'medical model', instead offering a psycho-social view, recovery focus, and theoretical basis for understanding clients' distress. Psychologists can provide, or facilitate the development of, a working formulation to guide the

team. My observations of Psychologists helping staff to 'make sense' of client presentations align well with this described role. Staff from other professions have also recognised and valued Psychologists' formulations and alternative perspectives (Wood et al., 2019). Sharing a formulation does not have to be a formal or one-off event. Clinical Psychologists participating in a study by Christofides et al. (2012) related an ongoing process of "chipping in", taking any opportunities to suggest hypotheses or alternative perspectives and gradually sharing psychological knowledge. They felt working in this way helped staff to understand and relate to clients, as well as demonstrating the work of Psychologists to the team. Clinicians from other disciplines value informal consultations that provide opportunities for feedback or new insights (Wood et al., 2019).

Clinical Psychologists have also described MDTs as a dedicated time to think and reflect, together making sense of clients' situations and ensuring everyone was working from the same understanding (Christofides et al., 2012). They noted the importance of acknowledging all team members' experience and expertise, and the potential role of Psychologists as facilitators. I have definitely seen Psychologists acting as facilitators, especially when other staff members were distressed. Collaborative practice helps provide consistency of care for clients, as well as encouraging creativity when team members feel they can discuss ideas freely and safely (Nutt & Keville, 2016). Psychologists can support the process by modelling effective peer consultation and reflective practice for other team members (Onyett, 2007). Supporting the team has been identified as a key role of Psychologists by other staff, who wanted Psychology to be a more visible and accessible part of inpatient mental health teams (Wood et al., 2019).

Reflections on Current and Future Practice

As a result of my research and placement experiences, I began my internship with an appreciation of the value of good MDTs and a willingness to take part. Encouraged by other staff, I have been fully involved in discussions, both asking questions and offering ideas. My acceptance as a Psychologist within the team has been both a pleasure and a privilege but this comes with some challenges when they appear to forget I am still only an *Intern* Psychologist. It can be difficult to balance my desire to contribute and awareness of my training with the knowledge that I still have little experience and a lot to learn. I have needed to remind staff that decisions about referrals must be made by my supervisors and practice checking myself before speaking, offering suggestions tentatively, and using a socratic style of questioning rather than making definite statements. This is, of course, good practice for interactions with clients too and I have noticed improvements in my ability to work collaboratively in a range of contexts.

I have benefitted from the development of supportive collegial relationships of my own and continue to learn from other staff each week. MDTs offer opportunities to not only receive advice for my own clients, but also to learn from the presentations of clients who would not normally be referred to Psychology. In my future practice, I want to continue to participate in MDTs with an open mind to learn from others, to share my knowledge so they can learn too, and to be conscious of my modelling of reflective collaborative practice. Researching this case study has increased my confidence in the value of good MDTs, a confidence I want to take into my future work with a willingness to advocate for this practice if necessary.

Supporting others as an integral part of the team aligns well with my own personal relational style and I have been encouraged by positive feedback from colleagues and the support for this approach in the literature. However, I have also experienced a student nurse taking my desire to be supportive as an opportunity for extended supervision, reminding me of the importance of maintaining professional boundaries and staying conscious of both my priorities and my scope of practice.

In a broader sense, my research gave me an understanding of the role of Psychology in contexts outside of typical mental health and the potential psychological needs of medical patients. Working in a hospital environment, I have seen Psychologists educating nurses about anxiety management strategies, advising on the support for patients experiencing psychosis in medical wards, and advocating for the assessment and treatment needs of patients with brain injuries. I want to be open to similar opportunities in my future practice with a willingness to think broadly and consult with other teams. Of course, this willingness will need to be balanced with clear priorities and an understanding of my role and responsibilities within my own service, ensuring I fulfil these first and check in with management or other team members when needed.

References

Almairac, F., Herbet, G., Moritz-Gasser, S., de Champfleur, N. M., & Duffau, H. (2015). The left inferior fronto-occipital fasciculus subserves language semantics: A multilevel lesion study. *Brain Structure and Function*, *220*(4), 1983–1995. https://doi.org/10.1007/s00429-014-0773-1

Bajada, C. J., Lambon Ralph, M. A., & Cloutman, L. L. (2015). Transport for language south of the Sylvian fissure: The routes and history of the main tracts and stations in the ventral language network. *Cortex*, *69*, 141–151. https://doi.org/10.1016/j.cortex.2015.05.011

Callahan, B. L., Macoir, J., Hudon, C., Bier, N., Chouinard, N., Cossette-Harvey, M., Daigle, N., Fradette, C., Gagnon, L., & Potvin, O. (2010). Normative Data for the Pyramids and Palm Trees Test in the Quebec-French Population. *Archives of Clinical Neuropsychology*, *25*(3), 212–217. https://doi.org/10.1093/arclin/acq013

Chang, W.-H., Pei, Y.-C., Wei, K.-C., Chao, Y.-P., Chen, M.-H., Yeh, H.-A., Jaw, F.-S., & Chen, P.-Y. (2018). Intraoperative linguistic performance during awake brain surgery predicts postoperative linguistic deficits. *Journal of Neuro-Oncology*, *139*(1), 215–223. https://doi.org/10.1007/s11060-018-2863-z

Chen, L., Lambon Ralph, M. A., & Rogers, T. T. (2017). A unified model of human semantic knowledge and its disorders. *Nature Human Behaviour*, *1*(3), 0039. https://doi.org/10.1038/s41562-016-0039

Christofides, S., Johnstone, L., & Musa, M. (2012). 'Chipping in': Clinical psychologists' descriptions of their use of formulation in multidisciplinary team working. *Psychology*

and Psychotherapy: Theory, Research and Practice, *85*(4), 424–435. https://doi.org/10.1111/j.2044-8341.2011.02041.x

Conner, A. K., Briggs, R. G., Sali, G., Rahimi, M., Baker, C. M., Burks, J. D., Glenn, C. A., Battiste, J. D., & Sughrue, M. E. (2018). A Connectomic Atlas of the Human Cerebrum—Chapter 13: Tractographic Description of the Inferior Fronto-Occipital Fasciculus. *Operative Neurosurgery*, *15*(suppl_1), S436–S443. https://doi.org/10.1093/ons/opy267

De Witt Hamer, P., Gil-Robles, S., Zwinderman, A., Duffau, H., & Berger, M. (2012). Impact of Intraoperative Stimulation Brain Mapping on Glioma Surgery Outcome: A Meta-Analysis. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology*, *30*, 2559–2565. https://doi.org/10.1200/JCO.2011.38.4818

Gamboz, N., Coluccia, E., Iavarone, A., & Brandimonte, M. A. (2009). Normative data for the Pyramids and Palm Trees Test in the elderly Italian population. *Neurological Sciences*, *30*(6), 453–458. https://doi.org/10.1007/s10072-009-0130-y

Gernsback, J. E., Kolcun, J. P. G., Starke, R. M., Ivan, M. E., & Komotar, R. J. (2018). Who Needs Sleep? An Analysis of Patient Tolerance in Awake Craniotomy. *World Neurosurgery*, *118*, e842–e848. https://doi.org/10.1016/j.wneu.2018.07.064

Gudayol-Ferré, E., Lara, J. P., Herrera-Guzman, I., Böhm, P., Rodés, E., Ansaldo, A. I., & Peña-Casanova, J. (2008). Semantic memory as assessed by the Pyramids and Palm Trees Test: The impact of sociodemographic factors in a Spanish-speaking population. *Journal of the International Neuropsychological Society*, *14*(01). https://doi.org/10.1017/S1355617708080168

Han, Z., Ma, Y., Gong, G., He, Y., Caramazza, A., & Bi, Y. (2013). White matter structural connectivity underlying semantic processing: Evidence from brain damaged patients. *Brain*, *136*(10), 2952–2965. https://doi.org/10.1093/brain/awt205

Klein, L. A., & Buchanan, J. A. (2009). Psychometric properties of the Pyramids and Palm Trees Test. *Journal of Clinical and Experimental Neuropsychology*, *31*(7), 803–808. https://doi.org/10.1080/13803390802508926

Lemée, J.-M., Bernard, F., Ter Minassian, A., & Menei, P. (2018). Right Hemisphere Cognitive Functions: From Clinical and Anatomical Bases to Brain Mapping During Awake Craniotomy. Part II: Neuropsychological Tasks and Brain Mapping. *World Neurosurgery*, *118*, 360–367. https://doi.org/10.1016/j.wneu.2018.07.099

Moritz-Gasser, S., Herbet, G., & Duffau, H. (2013). Mapping the connectivity underlying multimodal (verbal and non-verbal) semantic processing: A brain electrostimulation study. *Neuropsychologia*, *51*(10), 1814–1822. https://doi.org/10.1016/j.neuropsychologia.2013.06.007

Nitta, T., & Sato, K. (1995). Prognostic implications of the extent of surgical resection in patients with intracranial malignant gliomas. *Cancer*, *75*(11), 2727–2731. https://doi.org/10.1002/1097-0142(19950601)75:11<2727::AID-CNCR2820751115>3.0.CO;2-H

Nutt, K., & Keville, S. (2016). '… you kind of frantically go from one thing to the next and there isn't any time for thinking any more': A reflection on the impact of organisational change on relatedness in multidisciplinary teams. *Reflective Practice*, *17*(2), 221–232. https://doi.org/10.1080/14623943.2016.1159552

Olivarez, A., & Boroda, A. D. (2007). Review of The Pyramids and Palm Trees Test. In K. F. Geisinger, R. A. Spies, J. F. Carlson, & B. S. Plake (Eds.), *The seventeenth mental measurements yearbook*.

Onyett, S. (2007). *Working Psychologically in Teams* (New Ways of Working for Applied Psychologists in Health and Social Care). The British Psychological Society. https://www.wiltshirepsychology.co.uk/Working%20Psychologically%20in%20Teams.pdf

Talacchi, A., Santini, B., Casartelli, M., Monti, A., Capasso, R., & Miceli, G. (2013). Awake surgery between art and science. Part II: Language and cognitive mapping. *Functional Neurology*, *28*(3), 223–239.

Wood, L., Williams, C., Billings, J., & Johnson, S. (2019). The role of psychology in a multidisciplinary psychiatric inpatient setting: Perspective from the multidisciplinary team. *Psychology and Psychotherapy: Theory, Research and Practice*, *92*(4), 554–564. https://doi.org/10.1111/papt.12199