CENTRE FOR DECISION RESEARCH & EXPERIMENTAL ECONOMICS

University of Nottingham
UK | CHINA | MALAYSIA

Discussion Paper No. 2025-01

Malte Baader,
Urs Fischbacher,
Chris Starmer and
Fabio Tufano

March 2025

Measuring Creativity: Associative Thinking in Semantic Networks

CeDEx Discussion Paper Series
ISSN 1749 - 3293

CENTRE FOR DECISION RESEARCH & EXPERIMENTAL ECONOMICS

The Centre for Decision Research and Experimental Economics was founded in 2000, and is based in the School of Economics at the University of Nottingham.

The focus for the Centre is research into individual and strategic decision-making using a combination of theoretical and experimental methods. On the theory side, members of the Centre investigate individual choice under uncertainty, cooperative and non-cooperative game theory, as well as theories of psychology, bounded rationality and evolutionary game theory. Members of the Centre have applied experimental methods in the fields of public economics, individual choice under risk and uncertainty, strategic interaction, and the performance of auctions, markets and other economic institutions. Much of the Centre's research involves collaborative projects with researchers from other departments in the UK and overseas.

Please visit http://www.nottingham.ac.uk/cedex for more information about the Centre or contact

Samantha Stapleford-Allen
Centre for Decision Research and Experimental Economics
School of Economics
University of Nottingham
University Park
Nottingham
NG7 2RD
Tel: +44 (0)115 74 86214
Samantha.Stapleford-Allen@nottingham.ac.uk

The full list of CeDEx Discussion Papers is available at

http://www.nottingham.ac.uk/cedex/publications/discussion-papers/index.aspx

Measuring Creativity: Associative Thinking in Semantic Networks

Malte Baader[†] Urs Fischbacher[‡] Chris Starmer[§] Fabio Tufano[¶]

March 31, 2025

Abstract

Identifying creative ability and its determinants is crucial in understanding artistic and innovative achievements. Previous work has shown that performance across established creativity tasks does not correlate within participants. A potential reason for this finding is that most creativity tasks lack well-defined performance criteria. In this paper, we develop a novel tool for measuring creative ability and assess its performance through experimental tests. We construct a semantic network serving as the underlying structure of our tool. Based on this network, participants perform two associative thinking tasks, *Local Search* and *Depth Search*. We characterise each task by relating it to an established measure of creativity, finding that performance in our proposed tasks is significantly related to their matched creativity task across several dimensions. Our new tool improves on established creativity tasks by utilising a predefined solution space. While capturing key features of established methodologies, it substantially increases on the ease of implementation and interpretation. In addition we also provide causal evidence on the effect of incentives on our tool.

Keywords: Creativity, Associative Thinking, Methodology

JEL Codes: C91, D03, O39

We are grateful to participants at the Konstanz innovation workshop 2024, ASFEE Conference 2019, ESA European Meeting 2019, JDMx Conference Trento 2019, CCC Conference 2019 and NIBS Workshop 2019 as well as all members of CeDEx at the University of Nottingham. We also thank Sandro Ambühl, Giuseppe Attanasi, Gary Charness, Werner Güth, Anna Hochleitner, Andrew Schotter and Martin Sefton for comments on this project. Financial support by the Network for Integrated Behavioural Sciences (ESRC - NIBS (RA1175-06)) and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC 2117-422037984 is gratefully acknowledged.

[†]Norwegian School of Economics (FAIR), NO; University of Zurich, CH, malte.baader@nhh.no.

[‡]University of Konstanz, DE; TWI, Kreuzlingen, CH, urs.fischbacher@uni-konstanz.de.

[§]University of Nottingham, UK, chris.starmer@nottingham.ac.uk.

[¶]University of Leicester, UK, fabio.tufano@leicester.ac.uk.

1 Introduction

A documentary about *The Beatles* released in 2021 depicts a young Paul McCartney sitting in a recording studio, composing their hit song '*Get Back*' out of seemingly thin air. Feats such as this trigger a natural fascination to understand creative achievements and to learn what makes individuals more or less creative. However, creativity is not limited to the arts, it also plays a vital role in other areas such economic innovation or scientific discoveries (e.g. Scott & Bruce, 1994; O'Connor & McDermott, 2004; Hittmár et al., 2014; Sarooghi et al., 2015). Thus, identifying a robust measure of creativity encompasses great opportunities for organisational and scientific applications.

Although several established measures exist, there are two particularly significant limitations in their usage. First, most existing measures treat creativity as a holistic concept, suggesting that individual creative ability spans across domains such as arts, language or music (Torrance, 1962; Guilford, 1967; Bradler et al., 2019). However, previous empirical evidence shows that there is no within-participant correlation across distinct creativity tasks considered holistic, questioning whether experimental tasks can in fact capture creativity spanning various domains. (Baader et al., 2022).

The second limitation concerns experimental implementation. One way to characterise creative tasks is by their degree of *openness* (Unsworth, 2001; Charness & Grieco, 2019). While in "*...closed creativity tasks, there is a specific and delineated goal*" (Charness & Grieco, 2019, p.2), open creativity tasks "*...represent unfettered thinking outside the box without any obvious underlying ex-ante goal or direction*" (Charness & Grieco, 2019, p.3). A limitation of current measures is that approaches used to build in openness to allow for creative expression have necessitated subjective scoring based on vaguely specified concepts of creativity. In addition to scoring not being objective, other participants or research assistants must be recruited as evaluators posing demands on research budgets.

In this paper, we address both limitations and develop a novel experimental tool to capture creative ability. First, our tool focuses on a single, relevant domain of creativity, *associative thinking*. Associative thinking is defined as "*the creative thinking process as the forming of associative elements into new combinations*" (Mednick, 1962, p.221). An illustrating example for its relevance is the 'Medici-Effect' in the 14^{th} century (Johansson, 2004). The wealth of the Medicis attracted artists, architects, scientists, philosophers and philanthropists who used associative thinking across disciplines to drive years of innovation and ultimately fostered the emergence of the Renaissance. The ability to associate insights to construct something novel thus becomes a key driver of entrepreneurial and scientific progress. This notion is also mirrored in a quote by Steve Jobs, co-founder of Apple, who described creativity as "*...just connecting things. [creative people] were able to connect experiences they've had and synthesize new things.*" (Jobs, 1996) This quote shows that for Jobs associative thinking is not only an ingredient of creativity but the essence of it.

Second, In addition our tool improves on methodological properties. Contrasting other creativity tasks, our tool includes an ex-ante defined solution space allowing for a sharp implementation and

objective performance measures. To define the solution space, we construct a semantic network consisting of 41,234 words with 275,247 links between them. Within this network, we ask individuals to perform two associative thinking tasks: First, *Local Search* where participants have to find as many associations to a presented word as possible; second, *Depth Search* in which participants travel through our network from a start word to a target word.

In addition, to test the properties of our two tasks, in a within-participant design, we benchmark each of them against a well-established creativity task. We find that behaviour in our tasks — as measured by performance and decision times — correlates significantly with performance in these established tasks, whilst also improving experimental properties such as objective scoring, individual implementation and ease of incentivisation.

Our paper contributes to several strands of the literature. First, our work directly contributes to the emerging experimental literature on creativity in economics (e.g. Eckartz et al., 2012; Erat & Gneezy, 2015; Bradler et al., 2019; Attanasi et al., 2019b,a; Charness & Grieco, 2019; Attanasi et al., 2020; Gneezy et al., 2021; Charness & Grieco, 2022, 2023, 2024). Reviewing the creativity literature in experimental economics Attanasi et al. (2021) identify various shortcomings of other commonly utilised tasks as outlined above. While addressing these shortcomings, our tool offers an easily implementable tool to conduct research on creativity, with a particular focus on associative thinking.

Second, our focus on associative thinking relates to an established literature in psychology (Mednick, 1962; Bowden et al., 2005; Batey & Furnham, 2006; Benedek et al., 2012; Beaty & Silvia, 2012; Lee & Therriault, 2013; Beaty et al., 2014; Verhaeghen et al., 2017) where most studies find associative thinking to play a significant part in creative performance. Furthermore, as we are exploring semantic associations, we are directly building on work by Gough (1976), who first introduced *word association tasks* as a measure of creativity. Subsequent research has extended his approach by examining distance of associations (Acar & Runco, 2014) or neural involvement in associative thinking (Whitman et al., 2010). We further contribute to this literature by providing experimental tasks that capture two distinct types of associative thinking within the same network environment.

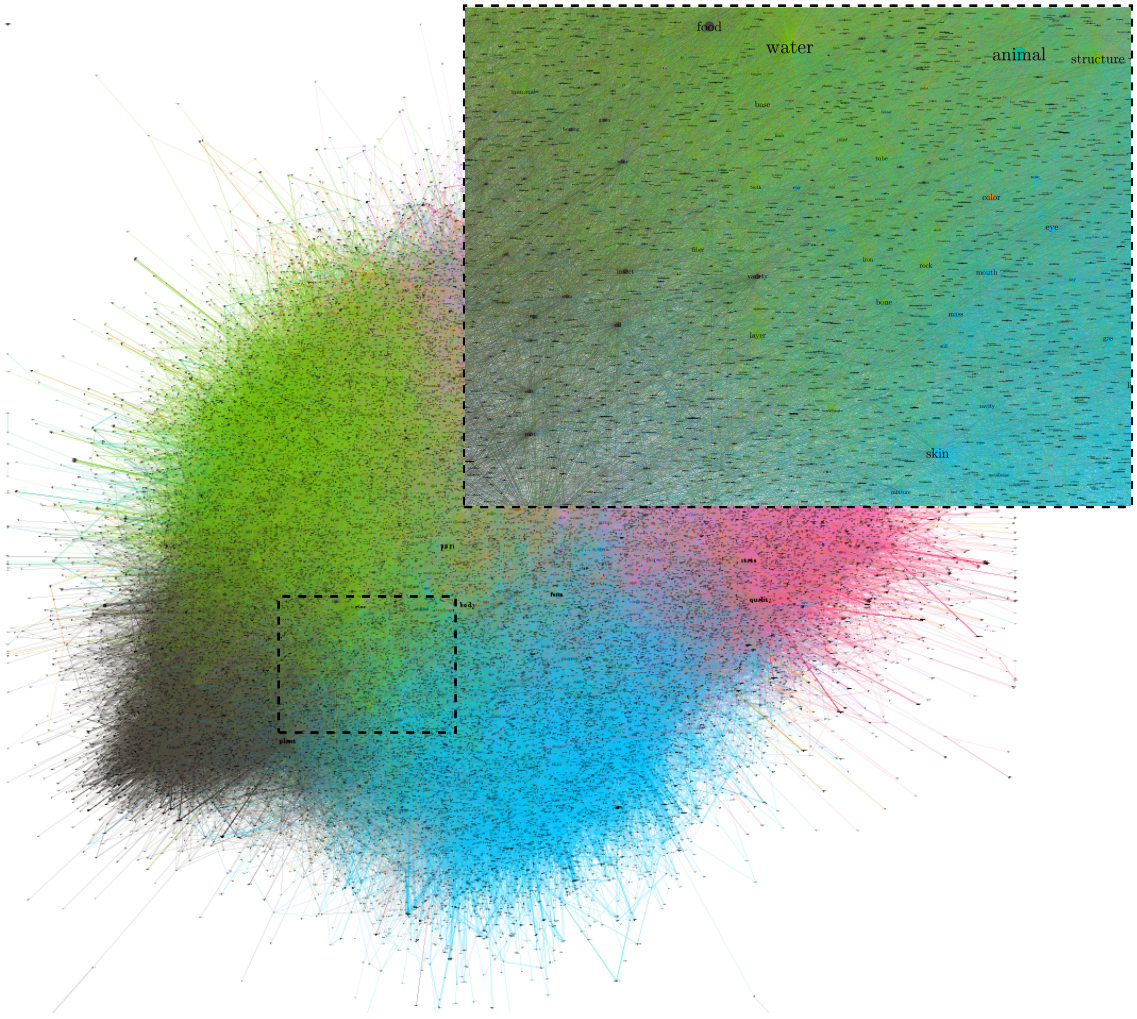
Third, we also examine incentive effects within our study and thus provide more evidence on the interaction of incentives and creative ability (Eckartz et al., 2012; Erat & Gneezy, 2015; Charness & Grieco, 2019; Bradler et al., 2019). In their highly influential paper, Charness & Grieco (2019) find incentives to only affect *closed* creativity tasks. However, we observe stronger effects of incentives in the *open* relative to the *closed* task. We attribute these differences to a non-monotonicity in how incentives affect creative performances across different levels of openness.

This paper is structured as follows. In Section 2, we outline the construction of our semantic network. In Section 3, we describe our two developed tasks. In Section 4, we provide the experimental design to compare them against existing measures of creativity. In Section 5, we present results. In Section 6, we discuss our findings and provide concluding remarks.

2 The semantic network

At the heart of our tool lies a semantic network, where words are connected based on their meaning. To connect words by meaning, we base our network on dictionary definitions. All words within the same dictionary definition share some relationship as they all refer to a unique underlying topic (i.e., the word explained). For example, the definition of procrastinator is "*someone who postpones work especially out of laziness or habitual carelessness*". When examining this definition, it is apparent that *nouns* deliver more meaningful information, *procrastinator*, *work*, *laziness*, and *carelessness*, whereas verbs, adjectives, or adverbs have less shared meaning. Therefore, within our semantic network, we define a link between two words as *any two nouns occurring in the same dictionary definition.* As the same noun is likely to occur in several definitions, we can then construct a network of all linked nouns

Figure 1: The semantic network



Note. Colours correspond to the 'modularity class', which is assigned by an algorithm identifying more connected words. Moreover, larger words represent more connected words. The image on the top-right captures the part of the network highlighted by the dashed rectangle.

in a dictionary. Using a *natural language processing tool*, we extract all nouns in the WordSet dictionary entailing around 177,000 entries, and identify their respective links.[1]

Overall, this yields a network spanning across 41,234 words with a total of 275,247 links between them.[2] The network encompasses all one-word nouns in the English language and their relationship with each other. Each word is on average connected to 13.44 other words and the diameter of the network, i.e. the longest possible path between two words, is 8, showing a considerable distance between words.[3] The word with the most links (*person*) has 4,296 other words connected to it, while there are 4,226 words (10.3%) that only have a single association. Figure 1 depicts the entire network, where coloured components are more interlinked and words with more links are larger.[4]

3 Developing two associative thinking tasks

Having constructed our network, we now use it as the underlying structure for two distinct associative thinking tasks, labelled *Local Search* and *Depth Search*.

In *Local Search*, participants are presented with a single word from the network and asked to list as many direct links as possible to the presented word within two minutes. In line with our network structure, participants are only allowed to submit singular, one word, English nouns and each submitted word is scored according to our predefined network. Participants receive one point for each correctly linked word but do not receive any feedback regarding the validity of their submissions during the task. This task assesses the ability to use associative thinking for divergent or open creativity, meaning the ability to generate diverse ideas by association from a single provided starting point (Guilford, 1950; Torrance, 1966; Guilford, 1967; Silvia et al., 2008; Palmiero et al., 2022). Due to the underlying structure, our semantic network provides a perfectly ex-ante defined solution space, ensuring that the task can be automatically scored without any additional, potentially subjective, evaluations by research assistants or other participants. To identify the final words presented to the participants, we follow a systematic four-step exclusion procedure outlined in Appendix C.1. This procedure leaves us with four final words for *Local Search* used in our experiment, which are *metal*, *tool*, *skin*, and *head*. Table 1 shows the number of associations for each word. A list of all valid associations for each *Local Search* word can be found in Appendix C.2. Figure 2a provides an example of the task and depicts all valid links to the word 'skin'.

In *Depth Search* participants have two minutes to travel from one word to another as fast as possible

[1] The NLP was developed by the Natural Language Processing (NLP) Group at the University of Stanford. See https://nlp.stanford.edu/software/lex-parser.shtml. The dictionary is a collaborative project initially based on the WordNet project of the University of Princeton.

[2] Note that not all dictionary entries are nouns, thus the number of words in the network is smaller than the number of entries. All details and files to recreate and utilise the network are available upon request.

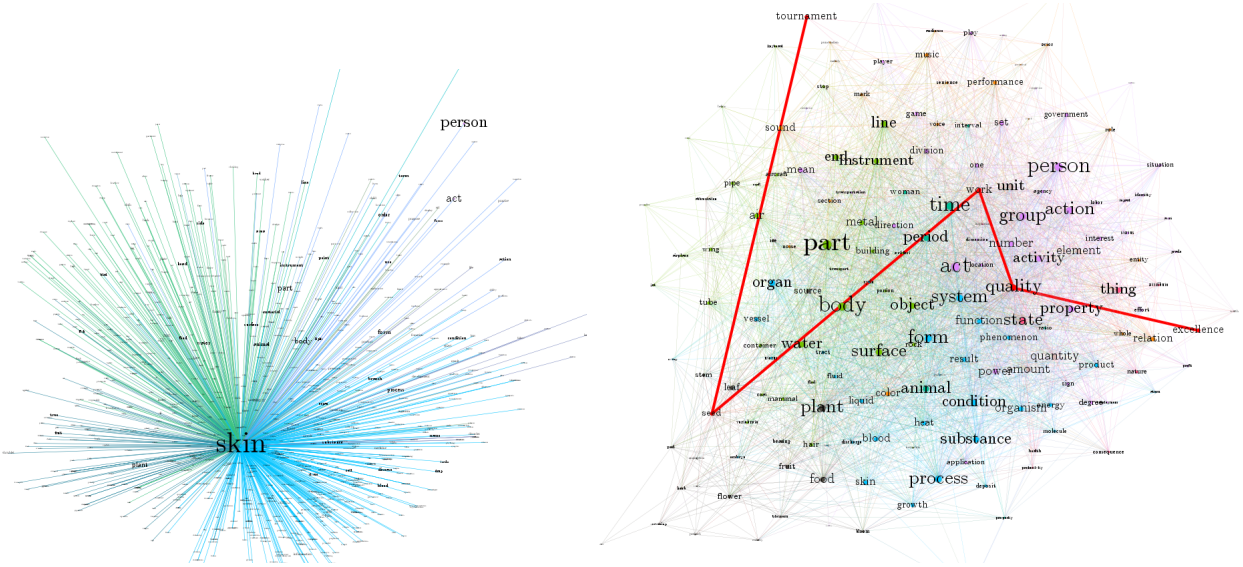
[3] For reference, according to the idea of 'six degrees of separation' by Milgram (1967), every pair of individuals living in the world is connected by a path of length 6.

[4] We use *Gephi* to illustrate networks, which includes a feature to compute a 'modularity class' to identify connected components of networks.

Figure 2: Visual representation of *Local Search* and *Depth Search*

a. Local Search

b. Depth Search



Note. Examples of *Local Search* (left) and *Depth Search* (right). *Local Search* shows all associations to 'skin', while *Depth Search* depicts a shortest path from 'excellence' to 'tournament'.

across the pre-defined network. Thus, participants have to consider deeper links within the network, including some that are multiple steps away from their current position. Moreover, to ensure that we disentangle *Local* and *Depth Search* as much as possible, we provide participants with all direct next steps once they reach a specific word. For example, starting at the word 'screwdriver', participants can observe all words associated to that word, for example 'handle', 'screw' or 'metal'. Once a participant selects the word 'handle', they can then observe all directly associated words for 'handle' and so on.[5] An example of *Depth Search* is provided in Figure 2b. It depicts the shortest path from 'excellence' (right) to 'tournament' (top). In our experiment, we then use *completion time* as our measure of performance.[6] Since the complete network entails words that are directly connected to more than 1,500 other words (e.g., *body*, *state* or *plant*), it is impossible to always visually present all links to participants. We thus do not use the entire network for *Depth Search* but construct smaller sub-networks. We provide all details on the construction of these sub-networks and our six-step exclusion process to identify relevant *Depth Search* paths in Appendix C.1. This process leaves us with exactly four sub-networks and paths that we test within our study.

Having described both of our tasks, we can also categorise them in terms of *task openness*, as outlined by Charness & Grieco (2019). Contrasting them, both of our tasks include a delineated goal and should therefore not be considered as entirely *open*. However, as *openness* is a continuous dimension,

[5] See screenshots of the task in Appendix B.

[6] We decided to use completion time and not the number of moves as a metric, as this allows participants to experiment during the task and move back after identifying a wrongly taken step. As we consider experimentation a key part of creative ability we did not want to penalise a wrongly taken step.

we still classify *Local Search* as substantially more open than *Depth Search*. Whilst participants know about the network, they do not know which exact answers are *valid* words. *Depth Search*, on the other hand, includes little ambiguity, which is why we classify it as *closed*.

4 Testing the two network tasks

We now explore our tasks' characteristics by relating them to established creativity tasks. Despite their shortcomings, we consider the wide utilisation of these established tasks to indicate a consensus of what is considered creative ability despite the challenges associated with their methodological limitations. If we find high correlations between measurements obtained using our new tools and those obtained using established ones, that strengthens the case for the validity and utility of our tasks. To validate *Local Search* we utilise the *Unusual Uses Task* (Guilford, 1967; Storm & Patel, 2014; Bradler et al., 2019; Dutcher & Rodet, 2022a,b; Stevenson et al., 2022), one of the most widely employed experimental creativity tasks in the academic literature. To validate *Depth Search* on the other hand, we benchmark it against the *Tower of London* (Shallice, 1982; Nitschke et al., 2017), an established tool for associative problem-solving. In the two sections we explain the connections between our two new tasks and their respective analogues in the established literature.

4.1 Local Search and Unusual Uses Task

For the *Unusual Uses Task*, participants are presented with an object and have to name as many *uses* for it as possible. In our experiment, we test four different objects, namely a *cardboard*, *tin can*, *paperclip* and *brick* and participants have two minutes to come up with uses for a given object. Upon completion, submissions are scored and examined for validity where each valid submission yields one point. Typically the *Unusual Uses Task* is assessed by three distinct metrics: the total number of uses submitted (*fluency*), the distinct categories covered (*flexibility*) and statistical scarcity (*originality*). Despite

Table 1: Problem overview of the *Local Search* task and the *Unusual Uses Task*

Local Search	# of possible associations
Metal	836
Head	979
Skin	831
Instrument	833
Unusual Uses Task	# of possible associations
Cardboard	undefined
Tin can	undefined
Paperclip	undefined
Brick	undefined

Note. A list of all possible solutions for the *Local Search* task can be found in Appendix C.2. The *Unusual Uses Task* does not include a defined number for associations due to its ill-defined solution space.

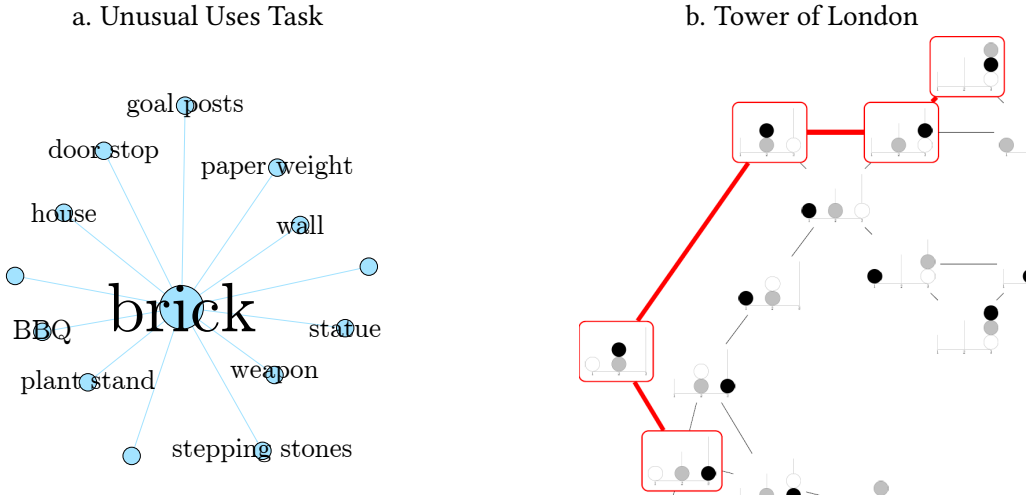
these dimensions being seemingly intuitive, the latter two are challenging to implement as they require external evaluators and ultimately rely substantially on which categories and uses are considered as distinct.[7] Since we wanted to include incentives within our study and due to previous work finding strong correlations between the three performance measures (Spearman's $\rho > 0.65$; p-values < 0.01), for this study we only use the dimension of *fluency* to evaluate performance in the *Unusual Uses Task* (Baader et al., 2022). Table 1 shows which objects we used for the *Unusual Uses Task* in the study and how the performance measure differs between tasks.

Despite these differences in performance measures, both tasks have many parallels and are intended to capture a similar aspect of creative thinking as both tasks address divergent thinking abilities. In other words, both focus on the ability of participants to generate many different solutions. Either task challenges participants to form associations from a predefined starting point and generate associated answers. To better illustrate this relationship, we provide schematic overviews of both tasks in Figures 2a and 3a.

4.2 Depth Search and Tower of London

To evaluate *Depth Search*, we compare it to another established and frequently used creativity task, the *Tower of London* (Krikorian et al., 1994; Berg & Byrd, 2002; Wagner et al., 2006; Köstering et al., 2015; Nitschke et al., 2017). In this task, participants have one minute to move three objects positioned on three rods as fast as possible from a start position to a target position. As in *Depth Search*, we use

Figure 3: Visual comparison of *Local Search* and *Unusual Uses Task*

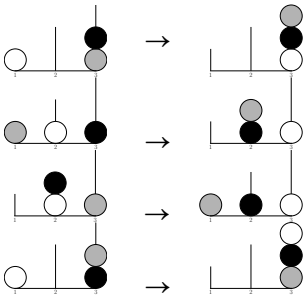


Note. Examples of *Unusual Uses Task* (left) and the *Tower of London* (right). The *Unusual Uses Task* indicates the undefined solution space by missing labels on the left. The example for the *Tower of London* shows a part of the overall solution network.

[7] Suppose the object provided is a *brick*. When providing answers, one participant could give uses as "building a house, building a school, building a church", whereas another person might just say 'building a building". Evaluators then have to decide whether the first participant submitted three uses, or a single use like the second participant. The *Unusual Uses Task* thus leaves substantial room for subjective interpretations and scoring, which complicates its implementation.

Table 2: Problem overview of the *Depth Search* task and the *Tower of London*

a. Depth Search	
Path	Shortest path length
dereliction → *ounce*	4
excellence → *tournament*	4
conduit → *outburst*	4
cold → *schematic*	4

b. Tower of London	
Path	Shortest path length
	6
	7
	6
	6

Note. Depth Search problems (panel a) include the seeds of the sub-networks in column 1. All *Tower of London* problems are within the same network. To select the problems for the *Tower of London* we only examine paths of length 6 or longer such that, similar to *Depth Search*, we have enough complexity to measure associative ability. We then randomly choose four out of the 213 possible paths, conditional on not being isomorphic. The networks can be found in Appendices C.3 and C.5.

completion time as our measure of performance. The three rods fit a different number of objects, ranging from one to three. Thus, participants must plan ahead to identify the best moves. Moreover, as there is a finite number of moves at each position, the solution space of the *Tower of London* can be displayed in a network, where solving a problem is equivalent to finding a path in that network.[8] For this reason, we consider the *Tower of London* to require similar reasoning as *Depth Search*. Again, to illustrate their similarities, we provide examples of both tasks in Figures 2b and 3b.

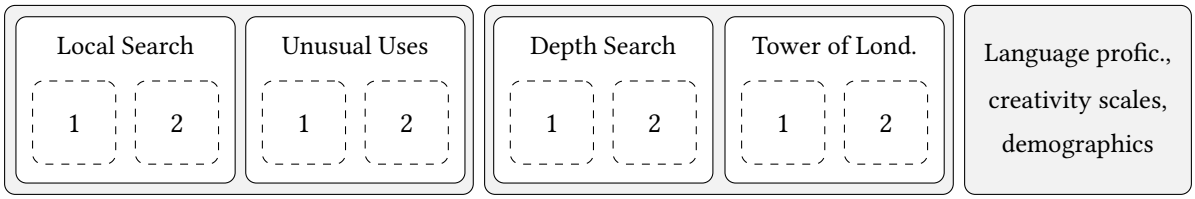
Table 2 lists the four problems for *Depth Search* and the *Tower of London*, including start and target words/positions as well as the shortest path length.

4.3 Design and Procedure

To test the validity of our two associative thinking tasks, we conduct a within-participant experiment where each person completes all four tasks described above. The two network tasks, *Local* and *Depth Search*, and both creativity tasks, the *Unusual Uses Task* and the *Tower of London*. Moreover, we ask participants to complete each task twice with randomly selected problems to study the tasks' robustness. Figure 4 provides an overview of our experimental design. To control for learning and order effects, we randomise the order of all four tasks and, within each task, of the two problems.

[8] See Appendix C.5 for a graphical representation of the network.

Figure 4: Experimental overview



Note. Each participant completes all parts outlined in the figure. The order of all tasks is randomised. The 2 problems for each task are randomly drawn from a set of 4.

We only recruited participants with English as their first language and elicited linguistic proficiency after completion of the tasks.[9] Further, we elicit self-reported associative ability (Jabri, 1991) and creativity (Runco et al., 2001) as well as collect demographic information. To explore incentive effects within our study, we vary incentives between participants. Half of the participants receive a flat payment for completing all tasks, whereas the other half receive piece-rate incentives. Participants could receive £0.50 for each valid word/use submitted in *Local Search* and *Unusual Uses Task*. During *Depth Search* (*Tower of London*), participants receive £0.10 (£0.20) for each second that was left from the total of two (one) minutes after reaching the target word.[10] At the end of the study, we randomly select one of the eight completed problems for payment.

The study was conducted online in April and July 2020. We recruited N = 400 participants using Prolific, 200 of which participated in the incentivised treatment and 200 in the non-incentivised one. The study was approved by the Nottingham School of Economics' Research Ethics Committee and programmed using LIONESS Lab (Giamattei et al., 2020). On average, participants took 30 minutes to complete the study and earned £12.50/hr.

Before presenting our results, we summarise our two main hypotheses. For each of our new associative thinking task, we expect a significant correlation with the corresponding creativity task, i.e. we expect correlation between *Local Search* and the *Unusual Uses Task* and correlation between *Depth Search* and the *Tower of London*. Moreover, in line with Charness & Grieco (2019), we expect incentives to have stronger effects on the more *closed* tasks, *Depth Search* and *Tower of London*.

5 Results

5.1 Local Search and Unusual Uses

We provide an overview of the most common words and uses submitted across problems in Appendix A, Figure A.1. We find, for instance, that the most common association for the word 'head' was 'hair', while the most common use mentioned for a 'paperclip' was using it as a 'paperclip' itself. Further, we

[9] See http://www.lextale.com/index.html. Numerous studies have shown that the task successfully captures individual language abilities.

[10] For instance, a participant taking 60 seconds to complete the task, earned (120-60)*£0.10 = £6.

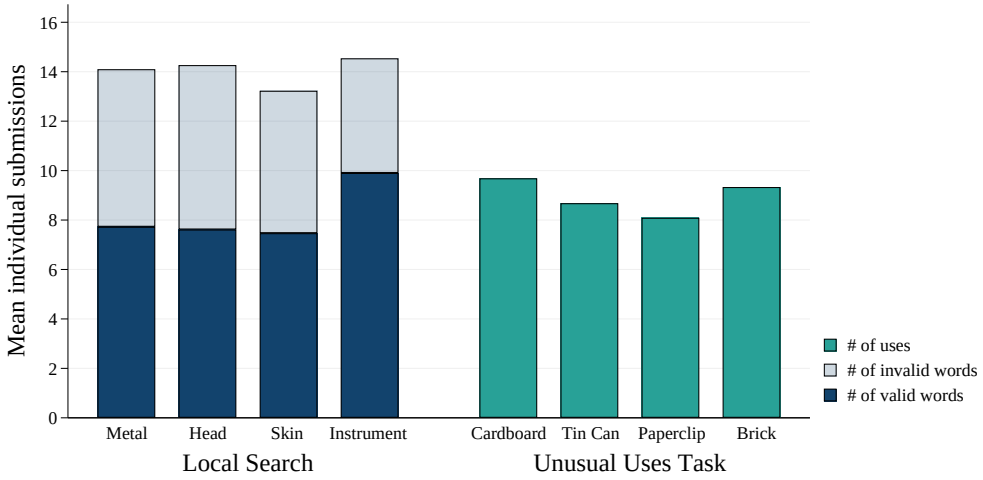
find that 58.22% of submitted words in *Local Search* were in fact linked in our network. When examining submissions in the *Unusual Uses Task*, we see that frequent submissions decline sharply after the first uses, which is also due to the fact that some uses have the same meaning (e.g., 'building' and 'build a house'). This exemplifies a key weakness of the *Unusual Uses Task*, where some participants submit similar uses as distinct ones, whereas others consider them identical.[11]

We next turn to our measure of performance in both tasks. Figure 5 shows the total number of submissions for *Local Search* (left) and *Unusual Uses Task* (right), including the separation into *valid* words for the former (dark).[12] We find strong similarities in average performance between both tasks (8.17 words and 9.27 uses respectively), suggesting a comparable degree of difficulty to generate new words/uses.

We regress performance in *Local Search* on performance in the *Unusual Uses Task* in Figure 6 and find a significant correlation between the tasks (Spearman's $\rho = 0.43$; p-value < 0.01). This suggests that *Local Search* does in fact measure the same associative ability that is captured by the *Unusual Uses Task*. This result is robust to the inclusion of controls ($\hat{\beta} = 0.51^{***}, SE = 0.07$). Our analyses further reveal that younger participants as well as those with higher linguistic proficiency perform better in *Local Search*. However, we do not find a relationship between gender, educational background or self-reported risk preferences and task performance (see Appendix A, Table A.1).

As participants completed two problems for each task, we can also apply the *Obviously Related*

Figure 5: Words/Uses across problems and tasks

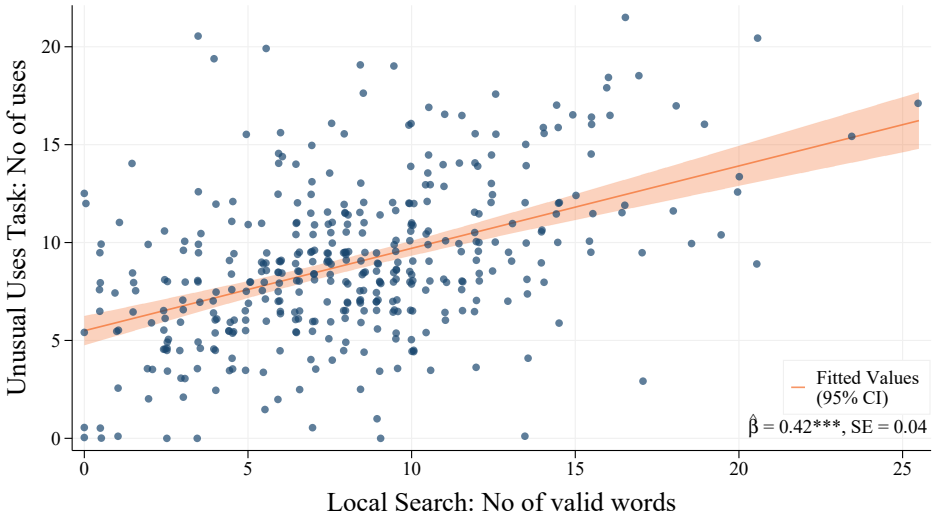


Note. The bar chart plots all submitted words and uses. For *Local Search* on the left, the bar is split between all submitted words (light) and words that are valid according to the underlying network (dark).

[11] The only option to resolve this issue is to evaluate all submissions individually and classify them as distinct or identical, which within our data corresponds to 5,243 evaluations and re-classifications. Evaluating submissions in *Local Search* by contrast can be directly implemented within the study, as only words linked in our network are valid.

[12] The number of invalid uses in the *Unusual Uses Task* is negligible and therefore not represented in Figure 5.

Figure 6: Relationship between the *Local Search* task and the *Unusual Uses Task*



Note. Points have been jittered slightly in both directions to reduce overplotting. Fitted line corresponds to a univariate regression between the *Local Search* (horizontal) task and the *Unusual Uses Task* (vertical). Regression coefficient and standard error provided in the bottom right corner of the graph.

Instrumental Variable (ORIV) approach (Gillen et al., 2019).[13] By instrumenting performances in both problems with each other, we remove potential measurement error to obtain *clean* correlations. Using ORIV, we find an increase in correlation (ORIV correlation = 0.59; p-value < 0.01), indicating that the relationship between the two tasks might be even stronger once accounting for noise. We interpret our result as evidence that *Local Search* and the *Unusual Uses Task* measure a 'common associative root', related to creative ability.

5.2 Depth Search and Tower of London

We now turn to the relationship between *Depth Search* and the *Tower of London*. Table 3 provides descriptive statistics of various performance metrics for each task. In addition to *completion time* as our main measure for performance, we also examine success rates and moves required. 70% (76%) of participants in *Depth Search* (*Tower of London*) successfully completed their paths within the given time limit. However, only 18% (24%) of participants in *Depth Search* (*Tower of London*), were able to find the shortest path. This is also reflected in the average number of moves in *Depth Search* (*Tower of London*) which is equal to 6.97 (9.04) and, thus, is above the minimum possible amount of four (six) moves.

We score all participants the maximum completion time if they did not succeed within the given time limit. As we expect all participants to eventually complete the problems, this scoring simply provides a lower bound for participants that failed to do so.[14]

[13] We find strong correlations between both problems that participants completed (Spearman's $\rho_{Unusual} = 0.73$; $\rho_{Local} = 0.50$, $p - values < 0.01$).

[14] All reported results also hold for restricting the sample to participants that finished the problem.

Table 3: Descriptive statistics - Depth Search and Tower of London

	Depth Search			Tower of London			
	Mean	Min	Max	Mean	Min	Max	N
Completion time	74.88	12	120	37.64	8	60	800
Success	0.70	0	1	0.76	0	1	800
Shortest path	0.18	0	1	0.24	0	1	800
# of Moves	6.97	4	39	9.04	6	23	609

Note. Depth Search is presented in columns 2-4. *Tower L* is in columns 5-7. The shortest path is conditional on successfully completing a problem. All of the *Depth Search* paths had a minimum of four moves. For the Tower of London they had a minimum path length of 6 or 7 (see Table 2).

To formally test whether performance is correlated, we regress completion time in *Depth Search* on completion time in the *Tower of London* in Table 4. We find a smaller but still statistically significant relationship between performances in both tasks (Spearman's $\rho = 0.21$; p-value < 0.01).[15] Again, this result is robust to the inclusion of additional controls (see Table 4), where, despite a reduction in effect

Table 4: Relationship between *Depth Search* and the *Tower of London*

	Completion Time	
	(3)	(4)
Tower of London	0.280***	0.156*
	(0.0941)	(0.0911)
Paths (Reference: dereliction → ounce)		
excellence → tournament	-5.564	-6.035
	(3.838)	(3.763)
conduit → outburst	58.23***	56.07***
	(4.153)	(4.103)
cold → schematic	-13.50***	-12.25***
	(3.734)	(3.685)
Demographics		
Age		0.386**
		(0.155)
Student		-5.922
		(4.262)
Language proficiency		-1.017***
		(0.218)
Order Controls	Yes	Yes
Additional Controls	No	Yes
Observations	800	760

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses.

Note. We provide results of Tobit regressions between completion times of *Depth Search* and the *Tower of London*. The regressions are censored at 120, the maximum time for *Depth Search*. Column (1) includes dummies for task order and controls for task comprehension. Additional controls in column (2) include gender, education, ethnicity, nationality, employment status, incentives and self-reported risk preferences.

[15] When excluding unsuccessful attempts, the correlation drops to Spearman's $\rho = 0.16$ (p-value < 0.01).

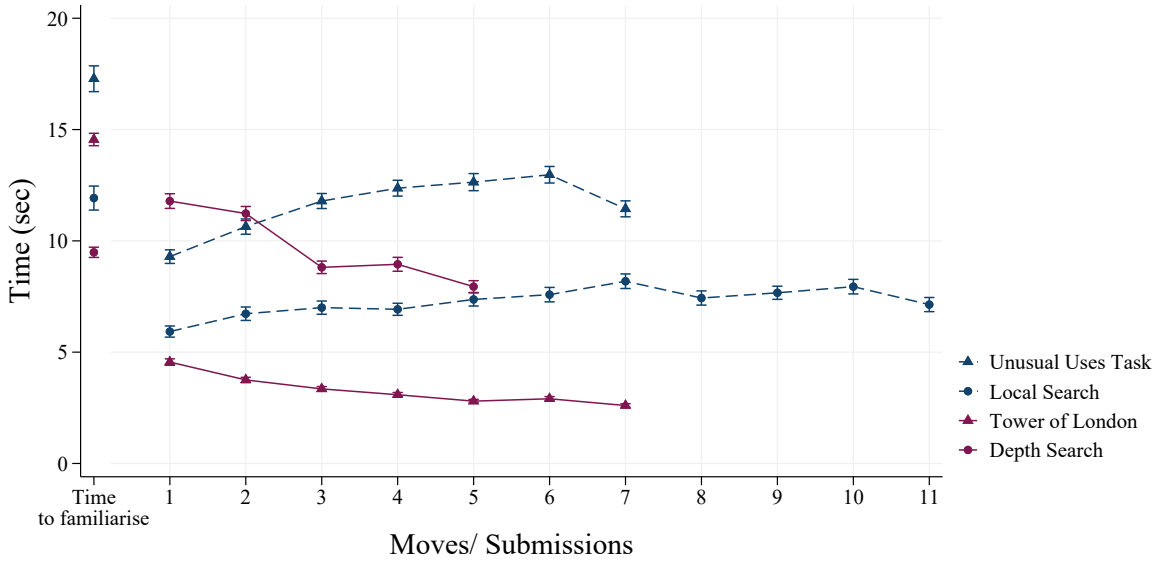
sizes, the relationship remains marginally significant (all p-values < 0.1). In addition, we again find that younger participants, and those with better language proficiency perform better in *Depth Search*.

After addressing measurement error using ORIV (Gillen et al., 2019), we find substantial increases in the correlation between *Depth Search* and the *Tower of London* (ORIV correlation = 0.55; p-value = 0.026). In fact, the *ORIV* coefficient is 2.5-times as large as the *raw* estimate, suggesting considerable measurement error in the individual elicitation. This result indicates that multiple measures of *Depth Search* are advisable as they provide significantly more information than a single measurement. We interpret our finding as evidence that *Depth Search* and the *Tower of London* capture a common form of associative thinking.

5.3 Differences across task types

As part of of our experimental setup, we also measured decision times across all four tasks. These reveal interesting correspondences in the decision time characteristics comparing across tasks. In particular, we measure the time for each individual to submit a new word/use in *Local Search* and the *Unusual Uses Task*, as well as the time it takes them to decide on a move in *Depth Search* and the *Tower of London*. Figure 7 illustrates how submissions and moves change over time in all four tasks, by capturing moves/submissions on the horizontal and decision times on the vertial axis. We can see that the times between submissions in *Local Search* and the *Unusual Uses Task* (blue) increase over time, whereas for *Depth Search* and the *Tower of London* (red) we observe a different pattern. The more participants approach the target word/position, the faster they make decisions about their next move. This evi-

Figure 7: Submission times across tasks



Note. Submission patterns for all tasks. Our set of tasks is represented by colour (blue and red), the two network and creativity tasks are indicated by different markers (circle and triangle). The first observation is disconnected as participants familiarised themselves with the interface. Observation numbers drop from left to right, as participants provided fewer uses/words or reached the target. At the final point of each line there are around 65% of the observations left.

dence complements our analysis from above and further highlights similarities in behaviour within the benchmarking tasks and differences between them.

Additionally, when we correlate performances across established creativity tasks as well as our proposed tasks, we find statistically significant correlations, which are larger across our proposed network tasks. (network tasks: Spearman's $\rho = 0.29$, p-value < 0.01; creativity tasks: Spearman's $\rho = 0.16$, p-value < 0.01). This does indicate that all four tasks require some common underlying ability.

Lastly, we also explore the relationship between experimental tasks and survey measures of creativity, finding no statistically significant relationship between self-reports and any of the four tasks (Appendix A, Table A.4)

5.4 The role of incentives

Despite the promising findings with respect to our proposed network tasks, one potential aspect that could explain correlation across all tasks in addition to underlying creative ability is motivation. If some participants have higher levels of intrinsic motivation and this motivation translates into performance, it could drive the statistical relationships we observed. To rule out this concern, we replicate all our results in the sub-samples with incentives, where motivation is increased (see Table 5). This suggests that motivation is not the the underlying reason for the significant relationships found above.

By studying incentive effects between participants, we can also speak more broadly to the literature of incentives for creativity. Examining the causal effect of incentives on creativity, we find that incentives only have a positive effect on performance in *Local Search* and the *Unusual Uses Task*. In both, participants submit significantly more valid answers when facing incentives (*Local Search* 7.79

Table 5: Benchmarking results across incentives

	No incentives		Incentives	
	(1) Local Search	(2) Depth Search	(3) Local Search	(4) Depth Search
Unusual Uses Task	0.400***		0.374***	
	(0.100)		(0.0791)	
Tower of London		0.240*		0.313**
		(0.142)		(0.140)
Age	-0.0473	-0.0411	-0.104***	0.185
	(0.0303)	(0.190)	(0.0277)	(0.180)
Female	-0.338	5.473	0.0862	4.368
	(0.790)	(4.263)	(0.611)	(3.622)
Additional Controls	Yes	Yes	Yes	Yes
Observations	195	195	185	185

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses.

Note. Individual performances are averaged across both rounds. Columns (1) and (3) provide linear regressions, whereas columns (2) and (4) present Tobit specifications. Additional controls in column (2) include gender, education, ethnicity, nationality, employment status, incentives and self-reported risk preferences.

to 8.54, $z = -2.15$, p-value = 0.03; *Unusual Uses Task* 8.40 to 9.47, $z = -3.50$, p-value < 0.01).[16] With respect to *Depth Search* and the *Tower of London*, by contrast, we fail to observe any significant effect of incentives (*Depth Search* 75.3 sec. to 74.5 sec., $z = 0.33$, p-value = 0.74; *Tower of London* 38.2 sec. to 37.1 sec., $z = 0.90$, p-value = 0.37). Provided the openness of our tasks, our findings suggest that incentive effects across levels of *openness* are non-monotonic. While, Charness & Grieco (2019) find that *open* creativity is not enhanced by incentives, also very *closed* tasks as *Depth Search* or the *Tower of London* do not respond to extrinsic incentives.

Finally, our finding that incentives affect task performance in the same way within the test pairs of tasks further supports our previous evidence that they indeed share similar characteristics.

6 Conclusion

In this paper, we develop and test novel tools to measure associative thinking. Associative thinking plays a key role in creativity, especially as a direct link to innovation. Our tool builds on an ex-ante defined solution space in form of a semantic network. This improves existing tasks by removing subjectivity when evaluating performances, allowing for simple incentivisation and focuses on a unique and relevant domain of creativity. Within this network, we propose two distinct associative thinking tasks: *Local Search* and *Depth Search*. While the former captures divergent thinking ability by generating as many *valid* links as possible, the latter requires the identification of *most relevant* links in order to reach a pre-defined target. We then test each of the tasks against established measures of creativity, the *Unusual Uses Task* and the *Tower of London*. We find strong associations between our new tasks and the corresponding benchmarking tasks. First, performance in each pair of tasks is significantly correlated. Secondly, we show that behaviour as identified through decision times is identical in each benchmarking pair of tasks. Lastly, we find that incentives affect each pair of tasks in a similar way. The accumulation of these factors supports our hypothesis that the same ability is required to perform well in *Local Search* and the *Unusual Uses Task* one the one hand as well as *Depth Search* and the *Tower of London* on the other. We therefore develop an tool that captures features of established creativity tasks whilst including objective performance measures in a clearly defined domain of creativity. Lastly, when exploring incentive effects across tasks our results suggest that a certain degree of task *closeness* might be required for positive incentive effects while incentives do not affect tasks that are very *closed*.

By formulating our tasks in the context of a semantic network, it is possible that the tools are thereby particularly tuned to measuring associative thinking connected to linguistic domains. We are aware of that concern, however, the tasks can be equally applied to other kinds of networks as well. Some examples are social networks among well-known individuals or economic networks between organisations. In theory, once an underlying network has been established, it can be transformed

[16] This positive incentive effect in the *Unusual Uses Task* is also in line with previous findings by Bradler et al. (2019).

into *Local* - and *Depth Search*. We encourage other researchers to compare such other domains with the semantic network we propose in this paper. Moreover, whilst our study focuses on the internal validity of the tasks, a next step could be to explore how they relate to creative achievements outside the experimental setting.

For these reasons, we consider our development just a first step to undertake more systematic creativity research in experimental economics. All in all, we provide an off-the shelf solution for those interested in studying creativity that can be applied in a variety of settings, from the laboratory to online studies or large-scale surveys.

References

Acar, S. & Runco, M. A. (2014). Assessing associative distance among ideas elicited by tests of divergent thinking. *Creativity Research Journal*, 26(2), 229–238.

Attanasi, G., Chessa, M., Ciucani, C., Gallen, S. G., et al. (2020). *Children's GrI-Creativity: Effects of Limited Resources in Creative Drawing*. Technical report, Groupe de REcherche en Droit, Economie, Gestion (GREDEG CNRS), Université C
ote d'Azur, France.

Attanasi, G., Chessa, M., & Llerena, P. (2021). A survey on experimental elicitation of creativity in economics.

Attanasi, G., Curci, Y., Llerena, P., Pinate, A. C., Ramos-Sosa, M. d. P., & Urso, G. (2019a). *Looking at creativity from east to west: risk taking and intrinsic motivation in socially and culturally diverse countries*. Technical report, Bureau d'Economie Théorique et Appliquée, UDS, Strasbourg.

Attanasi, G., Curci, Y., Llerena, P., Urso, G., et al. (2019b). *Intrinsic vs. extrinsic motivators on creative collaboration: The effect of sharing rewards*. Technical report, Bureau d'Economie Théorique et Appliquée, UDS, Strasbourg.

Baader, M., Fischbacher, U., Starmer, C., & Tufano, F. (2022). Investigating experimental measures of creativity. *Unpublished*. Available upon request.

Batey, M. & Furnham, A. (2006). Creativity, intelligence, and personality: A critical review of the scattered literature. *Genetic, Social, and General Psychology Monographs*, 132(4), 355–429.

Beaty, R. E. & Silvia, P. J. (2012). Why do ideas get more creative across time? an executive interpretation of the serial order effect in divergent thinking tasks. *Psychology of Aesthetics, Creativity, and the Arts*, 6(4), 309.

Beaty, R. E., Silvia, P. J., Nusbaum, E. C., Jauk, E., & Benedek, M. (2014). The roles of associative and executive processes in creative cognition. *Memory & Cognition*, 42(7), 1186–1197.

Benedek, M., Könen, T., & Neubauer, A. C. (2012). Associative abilities underlying creativity. *Psychology of Aesthetics, Creativity, and the Arts*, 6(3), 273.

Berg, W. K. & Byrd, D. L. (2002). The tower of london spatial problem-solving task: Enhancing clinical and research implementation. *Journal of clinical and experimental neuropsychology*, 24(5), 586–604.

Bowden, E. M., Jung-Beeman, M., Fleck, J., & Kounios, J. (2005). New approaches to demystifying insight. *Trends in Cognitive Sciences*, 9(7), 322–328.

Bradler, C., Neckermann, S., & Warnke, A. J. (2019). Incentivizing creativity: A large-scale experiment with performance bonuses and gifts. *Journal of Labor Economics*, 37(3), 793–851.

Charness, G. & Grieco, D. (2019). Creativity and incentives. *Journal of the European Economic Association*, 17(2), 454–496.

Charness, G. & Grieco, D. (2022). Creativity and ambiguity tolerance. *Economics Letters*, 218, 110720.

Charness, G. & Grieco, D. (2023). Creativity and corporate culture. *The Economic Journal*, 133(653), 1846–1870.

Charness, G. & Grieco, D. (2024). Creativity and ai. *Available at SSRN 4686415*.

Dutcher, G. & Rodet, C. S. (2022a). Creative capital generation: The role of diverse teams, experience and communication. *Experience and Communication*.

Dutcher, G. & Rodet, C. S. (2022b). Which two heads are better than one? uncovering the positive effects of diversity in creative teams. *Journal of Economics & Management Strategy*, 31(4), 884–897.

Eckartz, K., Kirchkamp, O., & Schunk, D. (2012). How do incentives affect creativity? Jena Economic Research Papers, no. 068.

Erat, S. & Gneezy, U. (2015). Erratum to: Incentives for creativity. *Experimental Economics*, 18(4), 760–761.

Fimbel, E., Lauzon, S., & Rainville, C. (2009). Performance of humans vs. exploration algorithms on the tower of london test. *PloS one*, 4(9), e7263.

Giamattei, M., Yahosseini, K. S., Gächter, S., & Molleman, L. (2020). Lioness lab: a free web-based platform for conducting interactive experiments online. *Journal of the Economic Science Association*, 6(1), 95–111.

Gillen, B., Snowberg, E., & Yariv, L. (2019). Experimenting with measurement error: Techniques with applications to the caltech cohort study. *Journal of Political Economy*, 127(4), 1826–1863.

Gneezy, U., Laske, K., & Schröder, M. (2021). Teams do not outperform individuals in a simple creative task. *Applied Economics Letters*, (pp. 1–6).

Gough, H. G. (1976). Studying creativity by means of word association tests. *Journal of Applied Psychology*, 61(3), 348.

Guilford, J. P. (1950). Creativity. *American psychologist*, 5(9), 444–454.

Guilford, J. P. (1967). *The nature of human intelligence*. New York, NY: McGraw-Hill.

Hittmár, Š., Varmus, M., & Lendel, V. (2014). Proposal of model for effective implementation of innovation strategy to business. *Procedia-Social and Behavioral Sciences*, 109, 1194–1198.

Jabri, M. M. (1991). The development of conceptually independent subscales in the measurement of modes of problem solving. *Educational and Psychological Measurement*, 51(4), 975–983.

Jobs, S. (1996). Steve jobs: The next insanely great thing. Wired. Retrieved from https://www.wired.com/1996/02/jobs-2/.

Johansson, F. (2004). *The medici effect.* Penerbit Serambi.

Köstering, L., Nitschke, K., Schumacher, F. K., Weiller, C., & Kaller, C. P. (2015). Test–retest reliability of the tower of london planning task (tol-f). *Psychological Assessment*, 27(3), 925.

Krikorian, R., Bartok, J., & Gay, N. (1994). Tower of london procedure: a standard method and developmental data. *Journal of clinical and Experimental Neuropsychology*, 16(6), 840–850.

Lee, C. S. & Therriault, D. J. (2013). The cognitive underpinnings of creative thought: A latent variable analysis exploring the roles of intelligence and working memory in three creative thinking processes. *Intelligence*, 41(5), 306–320.

Mednick, S. (1962). The associative basis of the creative process. *Psychological Review*, 69(3), 220.

Milgram, S. (1967). The small world problem. *Psychology Today*, 2(1), 60–67.

Nitschke, K., Köstering, L., Finkel, L., Weiller, C., & Kaller, C. P. (2017). A meta-analysis on the neural basis of planning: Activation likelihood estimation of functional brain imaging results in the tower of london task. *Human brain mapping*, 38(1), 396–413.

O'Connor, G. C. & McDermott, C. M. (2004). The human side of radical innovation. *Journal of Engineering and Technology Management*, 21(1-2), 11–30.

Palmiero, M., Fusi, G., Crepaldi, M., Borsa, V. M., & Rusconi, M. L. (2022). Divergent thinking and the core executive functions: A state-of-the-art review. *Cognitive Processing*, 23(3), 341–366.

Runco, M. A., Plucker, J. A., & Lim, W. (2001). Development and psychometric integrity of a measure of ideational behavior. *Creativity Research Journal*, 13(3-4), 393–400.

Sarooghi, H., Libaers, D., & Burkemper, A. (2015). Examining the relationship between creativity and innovation: A meta-analysis of organizational, cultural, and environmental factors. *Journal of Business Venturing*, 30(5), 714–731.

Scott, S. G. & Bruce, R. A. (1994). Determinants of innovative behavior: A path model of individual innovation in the workplace. *Academy of Management Journal*, 37(3), 580–607.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 298(1089), 199–209.

Silvia, P. J., Winterstein, B. P., Willse, J. T., Barona, C. M., Cram, J. T., Hess, K. I., Martinez, J. L., & Richard, C. A. (2008). Assessing creativity with divergent thinking tasks: Exploring the reliability and validity of new subjective scoring methods. *Psychology of Aesthetics, Creativity, and the Arts*, 2(2), 68.

Stevenson, C., Smal, I., Baas, M., Grasman, R., & van der Maas, H. (2022). Putting gpt-3's creativity to the (alternative uses) test. *arXiv preprint arXiv:2206.08932*.

Storm, B. C. & Patel, T. N. (2014). Forgetting as a consequence and enabler of creative thinking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(6), 1594.

Torrance, E. P. (1962). Non-test ways of identifying the creatively gifted. *Gifted Child Quarterly*, 6(3), 71–75.

Torrance, E. P. (1966). *Torrance Test of Creative Thinking: Norms Technical Manual*. Princeton, NJ: Research Edition Personnel Press Inc.

Unsworth, K. (2001). Unpacking creativity. *Academy of Management Review*, 26(2), 289–297.

Verhaeghen, P., Trani, A. N., & Aikman, S. N. (2017). On being found: How habitual patterns of thought influence creative interest, behavior, and ability. *Creativity Research Journal*, 29(1), 1–9.

Wagner, G., Koch, K., Reichenbach, J. R., Sauer, H., & Schlösser, R. G. (2006). The special involvement of the rostrolateral prefrontal cortex in planning abilities: an event-related fmri study with the tower of london paradigm. *Neuropsychologia*, 44(12), 2337–2347.

Whitman, R. D., Holcomb, E., & Zanes, J. (2010). Hemispheric collaboration in creative subjects: Cross-hemisphere priming in a lexical decision task. *Creativity Research Journal*, 22(2), 109–118.

Appendix for

Measuring Creativity: Associative Thinking in Semantic Networks

Malte Baader[1], Urs Fischbacher[2], Chris Starmer[3], Fabio Tufano[4]

[1]Norwegian School of Economics (FAIR), NO; University of Zurich, CH, malte.baader@nhh.no.

[2]University of Konstanz, DE; TWI, Kreuzlingen, CH, urs.fischbacher@uni-konstanz.de.

[3]University of Nottingham, UK, chris.starmer@nottingham.ac.uk.

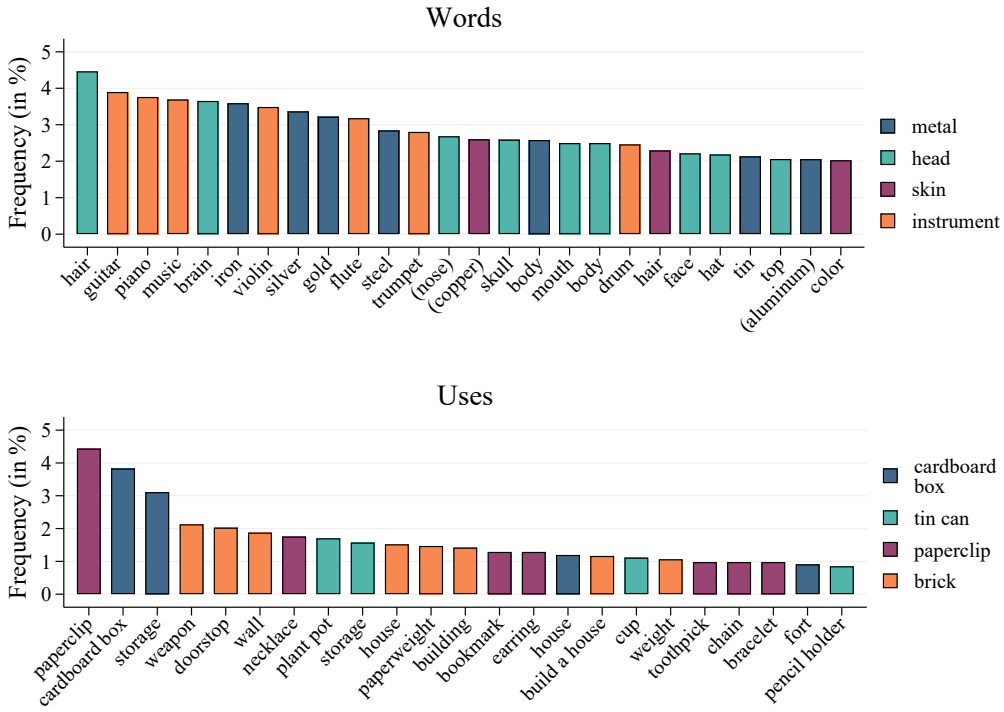
[4]University of Leicester, UK, fabio.tufano@leicester.ac.uk.

Contents

A Additional analyses

Table A.1 presents regression results of the relationship between the *Local Search* and the *Unusual Uses Task* (Models 1 and 2) and the Depth Search task and the Tower of London (Models 3 and 4). Model 1 has all submitted words as dependent, Model 2 only valid words. Model 3 uses Moves for the Depth Search task and ToL and Model 4 completion times.

Figure A.1: Most common submissions in the *Local Search* task and the *Unusual Uses Task*



Note. All words/uses submitted across all problems in the *Local Search* task and the *Unusual Uses Task*. Underlying words and objects are represented in different colours. The vertical axis shows the frequency of occurrence across all participants in percent. The words in brackets for the *Local Search* task indicate incorrect submissions.

Table A.1: Regression for comparisons within pairs of tasks

	LS & UUT		DS & ToL	
	(1)	(2)	(3)	(4)
# of Uses	0.509***	0.304***		
	(0.0723)	(0.0480)		
Tower of London			-0.0111	0.155*
			(0.0624)	(0.0910)
Incentives	1.163*	0.586	0.564	-1.790
	(0.615)	(0.385)	(0.538)	(3.156)
Additional Measures				
Vocabulary Test	0.100***	0.0952***	-0.0748***	-0.925***
	(0.0300)	(0.0196)	(0.0255)	(0.151)
Self-reported Risk	-0.157	-0.0721	0.159	1.020
	(0.142)	(0.0925)	(0.151)	(0.686)
Enjoy	0.510**	0.229*	-0.330**	-5.240***
	(0.199)	(0.132)	(0.153)	(0.959)
Difficulty	-0.736***	-0.412***	0.418***	5.041***
	(0.197)	(0.128)	(0.152)	(1.092)
Demographics				
Age	-0.127***	-0.0830***	-0.0287	0.384**
	(0.0259)	(0.0181)	(0.0261)	(0.155)
Student	-0.666	0.120	-0.844	-6.093
	(0.827)	(0.504)	(0.574)	(4.270)
Gender (Reference: Female, n = 230)				
Male (n = 166)	0.207	0.104	0.252	0.674
	(0.652)	(0.426)	(0.499)	(3.156)
Non-binary (n = 4)	6.530***	4.083***	-1.528	-7.982
	(2.095)	(1.072)	(2.433)	(20.45)
Education (Reference: College/A levels, n = 110)				
No formal qualifications (n = 8)	0.0135	-1.509	1.718	21.28
	(2.940)	(1.207)	(1.665)	(13.33)
Secondary school/GCSE (n = 48)	-0.409	-0.0251	-0.855	4.992
	(1.028)	(0.618)	(0.978)	(5.638)
Undergraduate degree (n = 158)	0.182	0.201	-1.218	-9.465**
	(0.721)	(0.466)	(0.784)	(3.844)
Graduate degree (n = 64)	0.439	0.137	-0.158	-7.817
	(0.880)	(0.570)	(0.911)	(4.806)
Doctorate degree (n = 8)	1.138	2.490	-1.750*	-13.34
	(2.368)	(1.662)	(0.952)	(9.005)
Order Controls	Yes	Yes	Yes	Yes
Additional Controls	Yes	No	Yes	Yes
Observations	759	759	760	760
R^2	0.427	0.380	0.195	-

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses.

Note. Model 1 has all submitted words as dependent, Model 2 only valid words. Model 3 uses Moves for the Depth Search task and ToL and Model 4 completion times. Additional controls: employment, nationality, problem dummies, order of tasks. and control question fails.

We present correlations for Local Search and Unusual Uses Task across rounds in Table A.2. The correlation coefficient presented is the relationship between first and second execution of the task within-participant.

Table A.2: Within-subject correlation across rounds for Local Search and Unusual Uses Task

	Round 2	
	Local Search	*Unusual Uses Task*
Round 1	0.529***	0.750***
	(400)	(400)

* p < 0.1, ** p < 0.05, *** p < 0.01.

As above, we present correlations for Depth Search and the ToL across rounds in Table A.3. The correlations are somewhat smaller in Depth Search, indicating lower levels of robust behaviour across rounds. Howver, the coefficient is still significant.

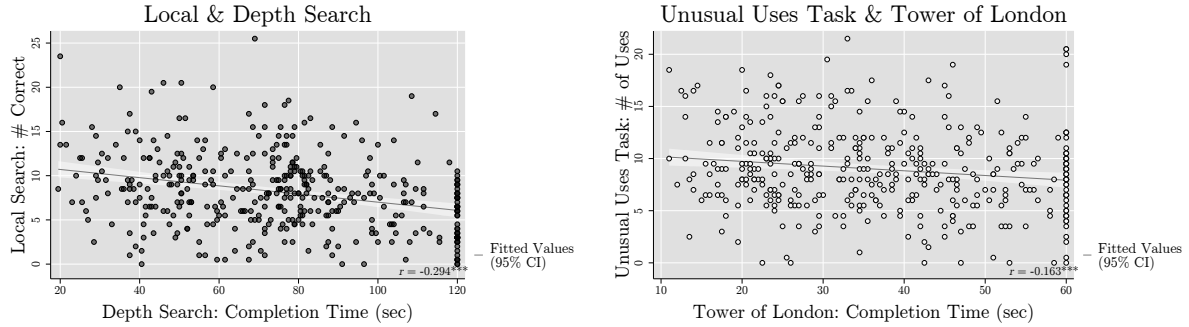
Table A.3: Within-subject correlation across rounds for Depth Search and ToL

	Round 2	
	Depth Search	*Tower of London*
Round 1	0.142***	0.471***
	(400)	(400)

* p < 0.1, ** p < 0.05, *** p < 0.01.

Table A.4 presents regression results for our collected survey measures. Both questionnaires are utilised with two factors as included in the table. We only find a marginal effect of the 'Ideational Behaviour Scale' with the Depth Search task, but no systematic relationship can be observed.

Figure A.2: Relationship within associative thinking and creativity tasks



Note. Scatter plots across sets of tasks. Network tasks on the left (*Local Search* vertical and *Depth Search* horizontal). Creativity tasks on the right (*UUT* vertical and *ToL* horizontal). Fitted line correspond to a univariate regression. Pearson's correlations are in the bottom right corners.

Table A.4: Regression survey measures

	Unusual Uses Task	Tower of London	Local Search	Depth Search
	(1)	(2)	(3)	(4)
Ideational Behaviour Scale (Runco et al. 2001)				
1st Factor	0.523	0.812	0.358	-6.784*
	(0.471)	(1.954)	(0.631)	(3.876)
2nd Factor	-0.175	0.297	0.0822	1.815
	(0.399)	(1.499)	(0.434)	(3.315)
Additional Controls	Yes	Yes	Yes	Yes
R^2	0.160	0.150	0.154	0.102
	(5)	(6)	(7)	(8)
Modes of Problem-Solving (Jabri, 1991)				
Associative	-0.0137	0.219	-0.286	2.675
	(0.247)	(1.113)	(0.293)	(2.379)
Bisociative	0.479	-0.313	0.121	-1.688
	(0.325)	(1.301)	(0.352)	(2.419)
Additional Controls	Yes	Yes	Yes	Yes
R^2	0.165	0.149	0.154	0.0981
Observations	390	390	390	390

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses.

All previous results between tasks hold when only examining the incentivised treatment. Table A.5 presents correlations across all tasks only for the incentivised subjects.

Table A.5: Within-subject correlation for incentivised treatment

	# of Uses	ToL Time	# of Correct	Depth Time
Creativity Tasks				
# of Uses	1			
ToL Time	-0.237***	1		
Network Tasks				
# of Correct	0.468***	-0.352***	1	
Depth Time	-0.196***	0.183***	-0.247***	1

* p < 0.1, ** p < 0.05, *** p < 0.01.

B Experimental instructions

[All instructions below were seen by the subjects. Each subject saw two randomly drawn tasks, for brevity we only present one in this section. . Two dotted lines represent a new screen in the programme. Below we only present the incentivised treatment. The instructions can be easily altered to remove the incentives.

The order of the tasks was randomised, we present just one order below. In this example the order is *UUT*, *Local Search*, *ToL*, *Depth Search*.]

Welcome to this study

In this study we will ask you to complete several tasks. You will be paid a fee of **£2.50 plus a bonus payment** depending on your performances for completing this study.

The study consists of **4 parts and a questionnaire**.

At the end of the study **one performance will be randomly chosen** to determine your bonus payment.

Instructions for each task will be provided as you go along. Please read all instructions carefully, answer the comprehension questions and complete the associated task.

Please click continue to proceed.

.....

.....

Instructions - Part 1 [Unusual Uses Task]

Please read through the instructions below carefully.

This part consists of 2 rounds. In each round, you have 2 minutes to come up with as many diverse uses for an object as possible.

For example

Please list as many, different and unusual uses for a rubber tyre as you can think of. Do not restrict yourself to a specific size of a tyre. You can also list uses that require several tyres. Do not restrict yourself to uses you are familiar with, but think of as many new uses as possible.

When thinking of your answers try to come up with as many diverse uses as you can. Answers are valid when they are either practicable or when their realisation is at least vaguely conceivable.

In this task for each valid use you submit, you receive a bonus of 50 pence. Thus, the more valid uses you submit, the larger your bonus will be. This means that if you submit 24 valid uses you would earn a bonus of 24 times £0.50 which equals £12.

After each use, you must submit your answer. You can submit by clicking 'Enter' or by using the submit button. You cannot change a use once it is submitted.

Once the time is up you will automatically be moved on.

Comprehension Questions

In order to move on please answer the following questions:

1) What is the main goal of the task?

- To come up with as many uses as possible
- To come up with as different uses as possible
- To come up with as many,different and unusual uses as possible

2) How much time do you have to come up with uses?

- 5 minutes
- 3 minutes
- 2 minutes

[**Subjects complete 2x2 minutes of the UUT with different objects**]

Instructions - Part 2 [Local Search]

Please read through the instructions below carefully.

This part consists of 2 rounds. In each round you have 2 minutes to come up with as many word associations as possible.

A valid association is a word that is connected to a pre-defined network based on dictionary definitions: Whenever two nouns appear in the same definition they are connected.

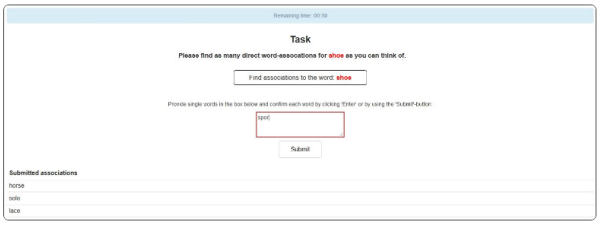
For each round, we will provide you with a specific word that you need to find associations to. Your task is to find as many words as possible that are connected to this specific word.

Please note that you can only enter words that are nouns (*What is a noun?*[**links to a definition of a noun**]) and singular. Examples are: mouse, colour or event. Terms with spaces (i.e. white noise) or nouns in plural are not valid. Please ensure the words are spelt correctly.

In this task for each correct association you submit you receive a bonus of 50 pence. Thus, the more correct associations you submit, the larger your bonus will be. This means that if you submit 24 correct associations you would earn a bonus of 24 times £0.50 which equals £12.

After each association, you must submit your answer. You can submit by pressing 'Enter' or by clicking the 'Submit' button. You cannot change an answer once it is submitted. You can find an example of the task screen below.

Once the time is up you will automatically be moved on.



Comprehension Questions

In order to move on please answer the following questions:

1) How much time do you have?

- 3 minutes
- 2 minutes
- 4 minutes

2) What is the main task?

- Copy words in a given time
- Do a mathematical calculation
- Provide as many associations to a word

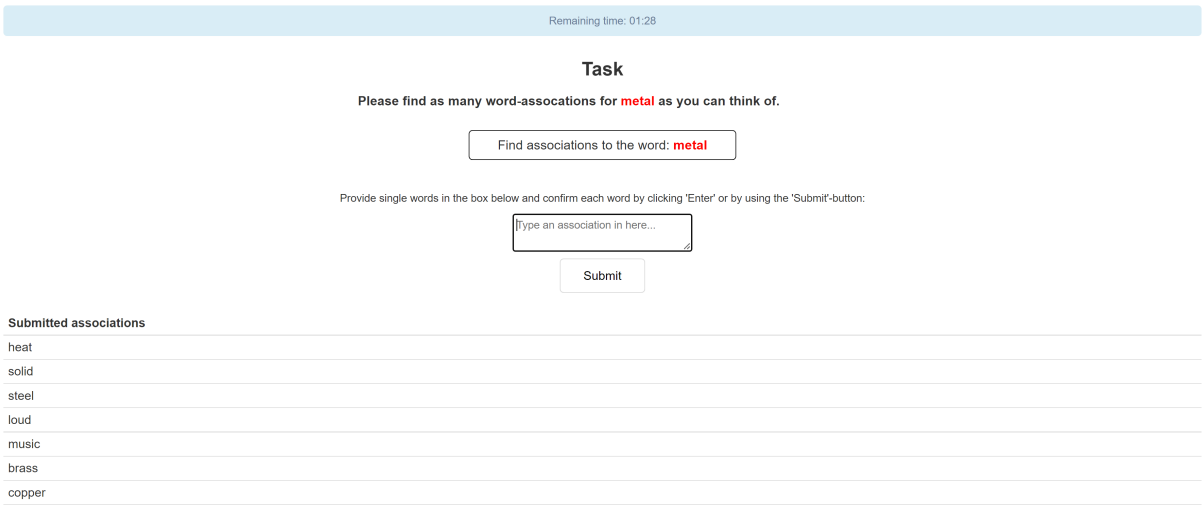
3) What is the role of the network?

- It does not have a special role
- It only provides a general direction of associations
- It defines valid associations

4) Which words can you enter?

- Any word can be entered
- Only nouns that are singular
- Only adjectives count as words

[**Subjects complete 2x2 minutes of the Local Search task with different words**]



Instructions - Part 3 [Tower of London]

Please read through the instructions below carefully

You are asked to play this part for 2 rounds. You have 60 seconds to complete each round.

On the screen there are 3 balls and 3 rods, which differ in length, 1 to 3, from left to right (see image below). The smallest rod can hold 1 ball, the middle can hold 2, and the largest rod can hold 3 balls at one time. In each round, you must move the 3 balls positioned on different rods to adjust a starting position to match an outcome position.

You can move the top ball from each rod to any of the other rods as long as the rod is not full. You can move the ball to any rod and you do not have to move the ball to a neighbouring rod.

Your task is to reach the outcome position **as fast as possible**.

In this task for each second left after you have finished, you receive a bonus of 20 pence. Thus, the faster you complete the task, the larger your bonus will be. So, imagine that you could solve the task in 0 seconds, then you would earn a bonus of 60 times £0.20 which equals £12.

Below you find an example of the task. You can move balls by clicking on the chosen ball and then the rod you wish to move it to. On the next page you will be able to practice the task.

Once you finish a round or run out of time you will be able to move on.

Comprehension Questions

In order to move on please answer the following questions:

1) What is the goal in each round?

- Reaching the outcome position as fast as possible
- Reaching the outcome position using the most moves
- I just have to reach the outcome position

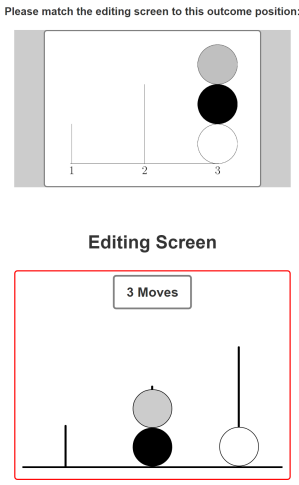
2) How much time do you have for each round?

- 30 seconds
- 120 seconds
- 60 seconds

3) Can you only move a ball to its directly neighbouring rod?

- Yes
- No

[**Subjects complete 1x practice and 2x60 seconds of the Tower of London with different paths**]



Instructions - Part 4 [Depth Search]

Please read through the instructions below carefully

This part consists of 2 rounds. In each round, you have 120 seconds to use connections to move from a start word to an end word. All connections are drawn from a pre-defined network based on dictionary definitions: Whenever two nouns appear in the same definition they are connected.

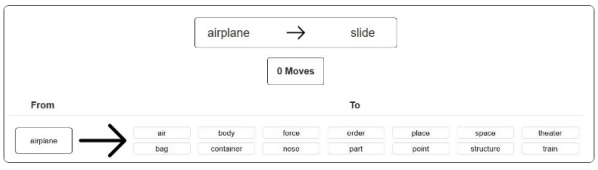
Your task is to reach the end word **as fast as possible.**

In this task for each second left after you have finished, you receive a bonus of 10 pence. Thus, the faster you complete the task, the larger your bonus will be. So, imagine that you could solve the task in 0 seconds, then you would earn a bonus of 120 times £0.10 which equals £12.

You are always provided with all possible next connections so your task is to find the connection that leads you towards the end word.

Below you find an example of the task screen. On the top of the screen you see the start and end word. Lower down the screen, on the left-hand side, you can see your current word (which in the example coincides with the start word) and on the right hand side all connections in alphabetical order. To make a move, just click on the connection you want to move to. On the next page you will be able to practice the task.

Once you finish a round or run out of time you will be able to move on.



Comprehension Questions

In order to move on please answer the following questions:

1) What is the main goal of the task?

- Avoid arriving at presented words
- Reach the end word as fast as possible

- Find the longest path between words

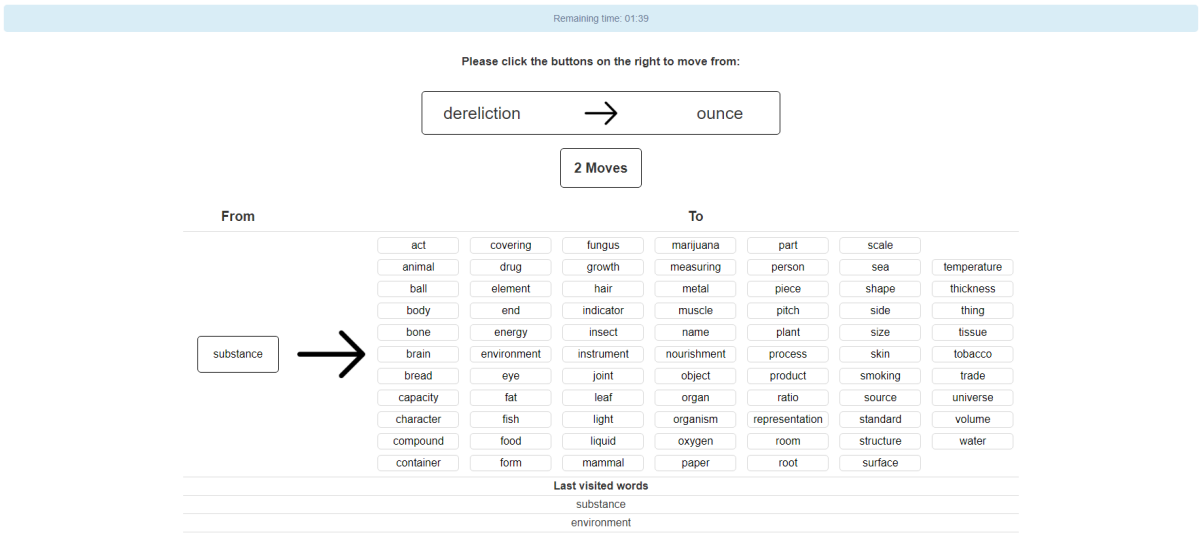
2) How much time do you have per problem?

- 90 seconds
- 120 seconds
- 60 seconds

3) How can you identify the next connections?

- I am provided with all direct connections
- I will have to type it in
- I am provided with a random selection of connections

[**Subjects complete 1x Practice and 2x120 seconds of the Depth Search task with different paths in distinct sub-networks**]



Questionnaire - Part 1/4 Please read through the instructions below carefully.

This questionnaire consists of about 60 trials, in each of which you will see a string of letters. Your task is to decide whether this is an existing English word or not. If you think it is an existing English word, click on “yes”, and if you think it is not an existing English word, click on “no”.

If you think that the word exists, even though you don’t know its exact meaning, you may still respond “yes”. But if you think this is not an existing word, you should respond “no”.

In this questionnaire, we use British English rather than American English spelling. For example: “realise” instead of “realize”; “colour” instead of “color”, and so on.

You can take as much time as you like for each decision but overall this part of the experiment should take about 5 minutes.

If everything is clear, you can now answer the comprehension question and then start the questionnaire.

Comprehension Questions

In order to move on please answer the following questions:

1) What is the main goal of the task?

- To spell check English words
- To say whether a word exists in English or not
- To create new English words

Questionnaire - Part 1/4

Please choose whether the presented word is an existing English word or not.
Click "yes" or "no" to make your choice.

Word:

unkempt

Yes No

Questionnaire - Part 2/4 Please answer each statement by indicating "**I am a person who enjoys...**" ranging from (**not at all**) to (**very much**). [**7-item Likert scale, 1 - Not at all; 7 - Very much**]

1) ...adhering to the commonly established rules of my area of work.
2) ...following well-trodden ways and generally accepts methods for solving problems.
3) ...being methodical and consistent in the way I tackle problems.
4) ...paying strict regard to the sequence of steps needed for the completion of a job.
5) ...adhering to the well-known techniques, methods and procedures of my area of work.
6) ...being strict on the production of results, as and when required.
7) ...accepting readily the usual and generally proven methods of solution.
8) ...being precise and exact about production of results and reports.
9) ...adhering carefully to the standards of my area of work.
10) ...being fully aware beforehand of the sequence of steps required in solving problems.
11) ...being confronted with a maze of ideas which may, or may not, lead me somewhere.
12) ...pursuing a problem, particularly if it takes me into areas I don't know much about.
13) ...linking ideas which stem from more than one area of investigation.
14) ...being fully occupied with what appear to be novel methods of solution.
15) ...making unusual connections about ideas even if they are trivial.
16) ...searching for novel approaches not required at the time.
17) ...struggling to make connections between apparently unrelated ideas.
18) ...spending time tracing relationships between disparate areas of work.
19) ...being 'caught up' by more than one concept, method or solution.

Questionnaire - Part 2/4 Please indicate for each statement how often from (never) to (very often) a statement occurs to you. [**7-item Likert scale, 1 - Not at all; 7 - Very much**]

1) I have many wild ideas.

2) I think about ideas more often than most people.

3) I often get excited by my own new ideas.

4) I come up with a lot of ideas or solutions to problems.

5) I come up with an idea or solution other people have never thought of.

6) I like to play around with ideas for the fun of it.

7) It is important to be able to think of bizarre and wild possibilities.

8) I would rate myself highly in being able to come up with ideas.

9) I have always been an active thinker — I have lots of ideas.

10) I enjoy having leeway in the things I do and room to make up my own mind.

11) My ideas are often considered "impractical" or even "wild."

12) I would take a college course which was based on original ideas.

13) I am able to think about things intensely for many hours.

14) Sometimes I get so interested in a new idea that I forget about other things that I should be doing.

15) I often have trouble sleeping at night, because so many ideas keep popping into my head.

16) When writing papers or talking to people, I often have trouble staying with one topic because I think of so many things to write or say.

17) I often find that one of my ideas has led me to other ideas that have led me to other ideas, and I end up with an idea and do not know where it came from.

18) Some people might think me scatterbrained or absentminded because I think about a variety of things at once.

19) I try to exercise my mind by thinking things through.

20) I am able to think up answers to problems that haven't already been figured out.

21) I am good at combining ideas in ways that others have not tried.

22) Friends ask me to help them think of ideas and solutions.

23) I have ideas about new inventions or about how to improve things.

Questionnaire - Part 2/4

1) Please tell me, in general, how willing or unwilling you are to take risks? Please use a scale from 0 to 10, where 0 means 'Completely unwilling to take risks' and a 10 means you are 'Very willing to take risks'. Pick a response from 0 to 10 to indicate below where you fall on this scale.[**11-item Likert scale, 0 - Completely unwilling to take risks; 7 - Very willing to take risks**]

1) What is your age? [**Numeric input**]

2) What is your gender? [**Text input**]

3) What is your ethnicity? [**Text input**]

4) What is your first language? [**Text input**]

5) What is the highest degree or level of school you have completed?

- No formal qualifications
- Secondary school/GCSE
- College/A levels
- Undergraduate degree (BA/BSc/other)

- Graduate degree (MA/MSc/MPhil/other)
- Doctorate degree (PhD/MD/other)
- Prefer not to say

[**For the four questions below. 7-item Likert scale, 0 - Not at all; 7 - Very much**]

How difficult did you find the part where you move from a start word to an end word?

Did you enjoy the part where you move from a start word to an end word?

How difficult did you find the part where you entered associations for a specific word?

Did you enjoy the part where you entered associations for a specific word?

.....

.....

[We then provide feedback about all performances and bonus payments. For brevity we do not present this screen in this appendix]

.....

.....

C Creativity tasks

C.1 Process to select final task problems

Local Search

Here we first describe how we selected the words participants faced during our study. Since any of the 41,234 words could serve as a base word for *Local Search*, we follow a systematic four-step procedure to choose words for the task:

1. We exclude all words that have fewer than 500 possible answers ensuring a large enough solution space (leaving us with 72 possible words).
2. We exclude any word that has a centrality score smaller than 0.30. In this instance we use 'eigenvalue centrality', which is a common network measure to evaluate *importance* of a node. Using it as an exclusion criteria consequently ensures that the remaining words are prominent and provide good starting points for participants (leaving us with 32 possible words).
3. We then exclude all words that are abstract nouns such as 'activity' or 'term' (leaving us with 10 possible words).
4. Lastly, we explore the overlap of solutions spaces and exclude words with more than 20% overlap (leaving us with six possible words).

From the remaining six words, we randomly select four that constitute our final words.

Depth Search

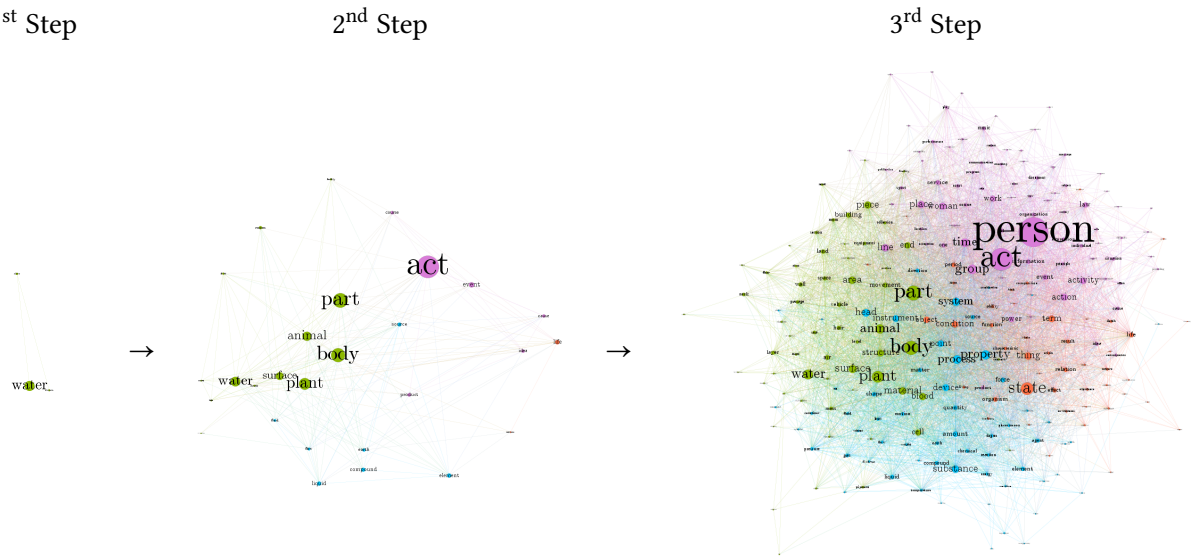
Before outlining the final path selection for *Depth Search*, we first describe how we construct sub-networks. For its construction, we start at a specific word, which we label the *seed* of the sub-network. We then identify all nouns that are in the *seed's* definition. Following that, we inspect the definitions of all these nouns to identify their nouns and iterate this step once more. This provides us with a set of words that all stem from the same *seed*.[17] To associate these words, we again rely on the links from our complete network. Figure C.1 visually represents the iterative process for the seed 'river'. The big advantage of these smaller sub-networks is that they allow us to visually present participants with all direct associations for *Depth Search*.

To identify which sub-networks to use in our experiment, we follow a similar process of exclusion as for the *Local Search* task.

1. We construct all possible 41,234 sub-networks.

[17] Note that it does not have to be the case that these words are in fact linked as in our complete network if 2 words during the iterative process never appear in the same definition but they do in some other definition not examined.

Figure C.1: Sub-network creation for *Depth Search* (seed: *River*)



Note. The 3-step iterative process of constructing a sub-network for the example of 'river'. In Step 1, we find the definition of 'river', in Step 2, the definitions of the nouns in the definition of Step 1. Step 3 follows the same procedure based on all definitions identified in Step 2.

2. we first exclude any sub-networks with fewer than 150 and more than 300 words (leaving us with 11,036 possible sub-networks).

3. We then randomly select 15 as candidates for our final tasks.

4. To ensure that there is sufficient distance between possible start and target words, we then only include sub-networks with a longest path larger or equal to 4 (leaving us with 6 possible sub-networks).

5. For the final paths in each sub-network for the experiment we only include paths of at least length 4.

6. We then exclude paths where either start or target word have fewer than three connections, to provide participants with sufficient possibilities and choices when starting a path or converging towards the target (leaving us with 49 possible paths across the six sub-networks)

From these we randomly select one path per sub-network and four of the six networks to again obtain four problems in total.[18]

[18]The final sub-network seeds and path details are presented in Table 2, for an illustration of the final four sub-networks, please see Appendix A, Figure C.2.

C.2 Local Search task

[Below we present all possible word associations in the *Local Search* task. For details on the words named by subjects please contact the authors.]

Word: metal

argent asteroid amount application animal action artifact article appearance attachment adornment air activity arrangement art affiliation armor armour alpenstock aglet aiglet ammunition area anvil argyle argyll apparatus appliance andiron armilla artillery awl armpit angle amalgam alchemist assayer analyst ash alloy asphalt bolt bone bent base bimetal bowl ball bullet belt box bonding bank basket bob blackjack basketball block battery babbitting bar bird bond ballpoint ballpen band beam backsaw bandsaw barrel baton bell bit bitumastic blade bicycle body bonnet bracelet boiler bullion brazier brasier brigandine buffer breech breechblock bushing brand brightness bloodstream badge bailiff bimetallism concrete cloisonne comal charcoal color current cutting center compound camp chelate chemical carbonyl ceramic conversion clip casting crutch cooking circle circuit chase cartwheel coating covering charge coupling construction combat competition cock coat cat carriage cage cosh cupel coop can club candle calk calkin canister cannister carbon cap casing cartridge caster carabiner chain chassis clapper clarinet circlet chimney cleat cleats coil column collet cockpit conductor component computer container corset cord crystal cramp crosshead cowl cringle cuff cutter cup cylinder cloth cowcatcher cowling chains cabasset coign coigne character currency coffee clay cavity cockfight crevice coal constable conductivity coin calcination chelation corrosion centimeter century device dash decoration dancing drink decalcomania design dog drum die disk disc diestock door dixie drawbar duct dulcimer drug drumstick doorknocker distance dimension diameter dancer dentist debaser degrader decalescence deterioration disease dross exchange enclosure extraction examination end echo element eyelet electrode edge enamel engine electroplate envelope emblem equipment earth elasticity electricity elixir engineer fluid form food firing flux filling frame foliation friction fatigue fabrication forging fire field fighting firedog fender fan fabric factory firearm fence ferrule fireplace fillet film fireguard fastener filter finger fingerboard flue fishplate flagon forge flashing foil front foundry framework furnace fret fretwork flour founder fume ferroconcrete game gold group gun golf galvanization galvanisation guard gamecock gas gong grommet grummet glass ground gridiron garment gaff gear gearbox globe glockenspiel gem grinder gurney gusset grate graver grid grill griddle grille grillwork guest gravity glaze gravy gunner gallon germanium garnet gemstone home hook handling holding head hanging hammer hand hammering horn hood horse hoop hearth handcuff hacksaw handlock helmet handsaw handspike heraldry heating heat hardware heel handle hunter honeycomb holder hole hilt hinge horseshoe hydrofoil hardness household hydrogen imitation iron interaction instrument incorporation interference image ironware ingot irons implement ivory ink ion inch impurity insulator joint junction jackstones jack jacket joist jewel jewelry jewellery kind key knucks knuckles karabiner keyring kickstand kern kettle knocker light living leading liquid lame locomotive ligature loop leg line lock lining lithography lead lid lamp lathe lamppost lattice latticework lacquer lacquerware leaf link ligament length log lever lunula lockbox letterpress lump lithium leather matte mass medium material mean metallic mineral mechanism mesh move mount mixture method music metalworking metalwork measuring miller magazine machine mail metalworks manacle match metalware missile mouth morion microwave motorcycle mirror mouthpiece manner mercury member metallurgy magnetohydrodynamics membership majorette melter metalhead metallurgist metalworker money tincture rock sun orbit staff point plastic sheath shoelace ribbon steel silver vessel water teeth test sheet net player weight string piece surface wood pen transfer paper reinforcement saw wheel rod sound rein weapon part wristwatch wrist pot ring track status rank unit standard value ratio strip monoxide page type time spoke vehicle wire tea worker product spring sailing rope striker side projection sole shoe pole thing object transport piston shackle trophy winner soldier wedge printer oxide result temperature process poisoning state oxidation shipping storage thread zither ornamental presence quality person phenomenon slowing rate scum pottery protection substance tooth work production stress screen plate workplace roof strength support preciousness superiority velocity tube use zinc rusting percussion utensil shell train set tool stretcher partition science pattern sphere saddle pommel vent smoke pipe watch shape necklace way one project printing reed saxophone stone post streetlight tumbler ornament strongbox valuable sulfide ore yarn place table technology plasma study wax refining solid monazite nail screw nut nugget phosphorus nitrogen nitride platinum osmium plumb oil piping press pressing perpendicular pick sugar purification timber upright position poker plating plastron pig pricket pin sheathing thickness shield pilot pan penicillamine patina setting pave plectrum piton plectron plexor pleximeter plessimeter pipage planchet plummet porringer round sonar ping plater potassium tin pewter quoin rail rapper radiator pair rings rule ruler refinement rubidium sport shot metal salt sculpture stabile stopping slug rigging spar screening shielding stay salamander scale projectile sap stove stovepipe smithy urn spigot samovar sand sandbox scarf scauper scorper silicon semiconductor scriber scribe shaper shoetree shim shear snips slat spline runner snow ski spike stake speculum rider stirrup straightness straightedge sword spread swage scrapheap sliver splinter workman textile shearer smith sodium strontium solder slag scoria toe tap tinker tongue slot token system tape plural tinsnips tapeline while sewing thimble tinsel thruster thermometer thermocouple voltage thermojunction trampoline triangle trivet trowel order tension turnbuckle temper toughness tarnish tip sign office tipstaff tapper troy upset xylophone resonator vibrato vibraphone vibraharp vibes pres-

sure welding path waveguide watchband watchstrap wristband screwdriver woodscrew workpiece weld windowpane wall window welder zarf

Word: skin

anterior adult arch attribute animal act air attempt arrangement affusion absence anaspid arthropod argasid armor ant aurochs antelope armadillo antenna abaya accelerator agal actuator arm appendage artery atrium attic ax axe arrow arrowhead aventail aura atmosphere aureole anteriority authority acetabulum angle artichoke army amphisbaena adviser ageratum andryala agueweed achillea acephalia acephalism acephaly asynclitism abscess alopecia back bolt bottom brachycephalic bone bowing base beetle bust boss bovid bull ball beanball beaner bullet basket beheading bob baseball basketball bow block baby bridle bird blackpoll beak baldpate bladdernose bulldog bluebottle bison bullhead bluehead boxfish branch balaclava band beam bark bedstead bearing basinet board boat body braid bollock bullfight blowhole blood brain bullethead bean bonce bemusement breed birth broccoli butter baldhead baldy brainworker butler bud borecole boneset banksia brodiaea beebalm bract balanitis balanoposthitis bighead baldness chicken capitate center composite cephalopod cetacean crustacean cause clout comb contraction covering charge crime class cephalometry ceremony carrot congregation crest cephalaspid crampfish cassowary carrion cockscomb coxcomb cockateel cockatiel clypeus calf cachalot collie corgi cat claw chlamyphore capuchin catfish checkrein club cabinet cam camail cap casque capote chanfron chamfron chest chessman chador chadar chaddar chuddar circlet clothing cloak coil college collider constituent computer container cord crook crown cushion cylinder cloth clubhead chapeau cervix cavity caul coxa chin condition concert cauliflower cabbage cole cos chicory clan cluster community colony country city comet cosmos chancellor capo chieftain chief colewort chrysanthemum cosmea catananche cornflower capeweed capitulum coin century clinocephaly clinocephalism cephalalgia cobalt calfskin dolichocephalic discharge device denial decapitation development decoration dilation demonstration danger dinosaur duck dog drum disk disc doll dome doornail drug drumstick defect distance diameter disbelief department datum divot diplomat dean don dahlia disease dengue evergreen empty execution expression elevation elongation examination end embryo eagle emperor elbow edge eye engine emblem ear equilibrium enarthrosis etiquette endive enemy empire educator edelweiss edema firm fluid fitting female flip feeling form feat formation foot feather flex friction fetometry foetometry fetus fowl foretop fin fish feeler frontstall face fighter fastener feature foliage foremast fusee fuzee front fruit forehead foreskin femur family foam froth father floret flower fleabane feverfew freshwater general group golf government gudgeon griffon gastropod gnu goat grenadier goujon grouper goby gas gurnard ground garment gaff gear governor grass garland glory grain gloriole genu growth gryphon griffin goldilocks gayfeather gosmore goldenrod groundsel goldfields high hip human hundred hook headshot header headshake headshaking headstand headship hat headlock water tail side part mouth shield pair insect robe toe wind hair kaffiyeh term tool handle tip mail hood neck light indication saint quality hipbone joint leaf heart plant serpent mythology sap herb native monster medicine labor presentation loss wool knee sign reverence submission shame hammer sculpture shoulder person victim inclination sac seal jaw mane horn sculpin spine male plate limb topmast opponent profession problem manservant servant household wine table tree shrublet seed inflammation top wrist shape metacarpus mollusk tentacle nail measurement skin organ order ray parrot whale spermaceti oil sperm ruff sheepdog pink monkey monk horse rein position state overcoat woman metal particle wreath victory vertex headdress weather organism rest vertebrate membrane lettuce root head salad university syndicate tribe weed pain muscle reaction turf piece skimming tropic orientation structure system sphere monarchy star platform tapering sucker lion shell ox sea mast spar sail panicle raceme shot hand photograph net soccer machine wagon help wrestling procession stream source viewer mass user individual pressure one school stimulation pus projection size height length liquid ship toilet representation thought pattern pole science passage text line inactivity matter subject movement pointer mark juncture role word percussion instrument kettledrum marimba ungulate protuberance shark hammerhead holocephalan holocephalian prey hooter hawk leather place hydra pike halberd helmet headpiece headrest kerchief headscarf headstall pillar stone statue herm headband panel headboard headgear harness hobnail halo standing surface headful information quantity margin hairline headache loaf sausage meat tongue pig headcheese savage trophy headhunter headman horseweed hortensia helianthus horsemint iron project ixodid point inion iceberg name region purple hue ironweed return vein jugular weight jerk scale jewfish triangle scarf knight kale kail queen king kingdom livelong lizardfish lid laurel loft meter stick match linstock stalk stem leopardbane time latency job minister ruler nation right monarch muzzle mudcat mallet sport masthead middle scalp mohawk thing metonymy organization title matriarchy matriarchate pond mill wheel millpond scorpion manticore mantichora manticora mantiger matriarch materfamilias mayor mayweed matchweed mistflower matchbush muskwood milkweed pendant molle megacephaly macrocephaly megalocephaly microcephalus microcephaly nod numbfish neb nailhead nimbus nucleus noodle noggin mind noddle napa nanocephaly overhead obeisance osteostracan owl ornithomimid ostrich notion reference truth reality sand missile rocket payload heat ogive obliquity skull occiput owlclaws orpin orpine spasm heel opisthotonos premier pick poll portfolio placoderm inch pachycephalosaur pachycephalosaurus reptile plesiosaur plesiosaurus puffbird pochard proturan mean spark plug pichiciago pichiciego pinche pickax pillow pickaxe pin pinhead rubber reflex plexor plessor percussor hole poncho sweater pullover puppet track positioner style ponytail precava pate protocol patriarchy patriarchate pantryman lance horseman picador patriarch paterfamilias officer president prexy prior pope pontiff pyrethrum umbel primrose primula platan phalacrosis queue iridescence paradise riflebird ringtail redhead weevil rostrum rattail object remora linen woolen sling rebozo teeth row rake rivet

variety radicchio romaine republic subshrub rabbitweed smash superior sovereign opening scapular somersault summersault summerset somersaulting sclaff thread shank screw phrase relation subordination shovelhead pectoral skate sauropod sticktight termite soldier snout snakefish suckerfish stargazer scorpaenoid scorpionfish slipover scapulary warmth secretary hub motion shaft setscrew humerus scapula shawl string snare skullcap snowball number sphinx spear spearhead spearpoint stature skeleton sinciput vertebra splenius seek savoy salsify scorzonera meeting summit rock supremacist skinhead stooper sunray sneezewort safflower succory snakeroot snakeweed sunflower sycamore priest tonsure member tack tape mechanism transport torpedo wing titmouse tomtit process wattle pheasant tragopan telsontail trunk toadfish thorax trunkfish testiere turban tabor tabour vessel topgallant tudung round thrift torso thoroughwort tidytips tansy toetoe teasel torticollis univalve undersecretary vulture ventail veil unit van vanguard hemisphere songbird waxwing wryneck wisent wildebeest wig witloof work power waterpower

Word: head

astringent avocado adult amphibian attribute abdominoplasty application animal action act appearance attack accumulation analysis acupuncture abnormality art absence anole arthropod antelope area araroba antiperspirant artery achromia achromasia account ability anus agnail aubergine apple aguacate anjou albino abrasion aspergillosis albinism algidity acne anemia acanthosis antifreeze back blond blonde bone buff blistering base break browning bite burn burning bag breakdown browse branding bird beak buffalo butterfish bonito branch blowfish band bark bearing bit border body brake bulla brand blood blemish birthmark blubber blackhead bile blister bleb birth blackheart blackamoor blastomycosis bedsore bruise bilirubin benzoin benjamin cut cosmetic chicken cherry color colour crack cancroid cell contact contusion cookery care circulation cosmetology cauterization cryosurgery cupping covering collection crest chimaera chameleon chamaeleon cobra coat cuticle calf cattle chamois coronet colugo chrysarobin canoe chest cephalexin cigar cigarette ciprofloxacin clothing coldcream container cream cup cloth cataplasm corn complexion callosity cheek clay comedo corium cartilage cortex canthus cranium clitoris cyst creepiness condition chafing crookneck cocozelle cherimoya cherimolla cataract child craftsman cumquat callus chromoblastomycosis coccidioidomycosis coccidiomycosis cancer carcinoma costiasis cyanosis cowpox cheilitis cellulitis cicatrix cicatrice clavus ceratin calfskin chammy concealer darkening demulcent distress development depilation dash deposit dilation damage dermabrasion diving doctor display dressing design dinosaur deodorant deodourant dermatome dapsone developer drug drip delicacy dryness depigmentation dewlap dermis derma dermatology degeneration dander dandruff dermatologist desquamation deterioration diaphoresis disorder dermatoglyphic disease dermatomycosis dermatophytosis dermatitis dermatosis dermatomyositis dracunculiasis dye exterior emollient escape exposure ending epilation enlargement elevation example erection ectoblast ectoderm exoderm edmontosaurus elephant eel eye embrocation effect elasticity exuviae epidermis eschar ear epicanthus eyelid epicranium extremity excoriation eruption eggplant exfoliation efflorescence epicarp exocarp ecchymosis ecdysis elastosis erythroderma erythema eczema erysipelas exanthem exanthema edema firm fluid fat fatty fell feeling form food failure force flare formation foot feather fold frame friction function fetus fleece filaria flea fin fish filefish face finger fruit focus fingerprint fissure flesh freckle follicle foreskin fingernail fingertip faculty fever fugu family flake fungus fermentation furuncle furunculosis formication freshwater gooseflesh goosebump germ goat garment gelatin grape gall gum gland genu growth grapefruit guava genip ginep goatskin gelatine glycerin glycerogelatin glycerogel hyaline human hybrid head horripilation hanging hand horn herpes hood hair hide hookworm hoof heart heat hangnail healing hidrosis hypothyroidism hypoadrenalism hypoadrenocorticism hickey hemorrhage hyperpigmentation hypopigmentation hives hyalin humectant hepatoflavin item interior impression identification instrument injection inflammation irritation injury inhalation insect implement infection integument intestine itch intertrigo infestation ichthyosis impetigo icterus injectant joint jacket jaundice juice kind kayak kumquat keratoderma keratoacanthoma keratodermia keratosis kraurosis keratin light liquid life location loss leap leg line liposuction layer lizard lip louse lappet lamb lemur limb lid liniment lotion lukewarmness luridness lividity lividness lentigo lung lymph lump loather lichen leprosy lesion lupus livedo leukoderma lack lambskin leather lactoflavin marine minute mass material mammalian mark medicine massage medication molter moulter mite mammal mole meloid mosquito makeup membrane mouth margin melanoderma muscle mummy macula macule milium melanocyte marrow melatonin mutation meal morello medlar mandarin mango mamoncillo molt molting moult moulting myxedema myxoedema molluscum melanosis melanism melanoma mange melanin name neck nail needle node nose nevus nerve nodule nectarine nettle neurofibromatosis neurofibroma neoplasm nitrogen orange opening organism obstruction order oxen ointment organ oil odor overgrowth overactivity onchocerciasis ovoflavin plum part patient protozoan person plant process pressure preparation prostration practice pattern petrissage puncture presence plethodont plug pachyderm pastern plectognath patch pit plate point product pigment pack peeler print plaster powder poultice purpose peach paleness pigmentation pallidness pallor posture physiology pore pulp palpebra perspiration pus plica prepuce palm problem pain pathology potato pumpkin pear peeling pericarp pox pityriasis pemphigus psoriasis prurigo papule pustule pimple pock petechia paresthesia paraesthesia pockmark pachyderma progeria pyrogallol protein parchment pigskin phenol propanediol quick round rash rose recognition reaction removal rock roughness response ray ranid reptile roundworm rhino rhinoceros race ridge rubefacient redness receptor result rind ringworm rhagades rosacea roseola rhinophyma resin riboflavin southeast sore stone surface substance spot solution salve snorkeling squash strip slice sweat surgery skincare stylostixis sting suction stimulation stinger shark salamander snake sarcoptid scale shell spine seed wrinkle stomach tip

treatment vegetable spore worker thickening vesicle thrust wound victim torture type whale ulcer wart tissue use tail tongue toenail vertebrate suede sheepskin side trade throat tract woman top toe shoe texture sole sensation soreness shrublet tree sputum udder scleroprotein tattoo scar skin toiletry scalp weakness worm water stiffness scab shedding vessel sunburn term symptom sheep stress weal vitamin weight thigh white tumor turkey temperature warmness wool secretion taste swelling variety teeth tone sickness vine smallpox swine wine table ungulate snout viscera vascularity snorkel touch sense site silurid soapfish sunscreen sunblock sign stigma scratch scrape sallowness stuffer sapodilla sapota satsuma slough scurf skinner sweating sudation scabies sporotrichosis steatocystoma seborrhea scleredema sun suntan shammy tan tanning toadfish triggerfish workplace tannery toner tepidness tepidity taxidermy tactility stimulus topognosis topognosia tangerine tangelo tanner taxidermist vapor transpiration typhus tinea undergarment underwear underclothes underclothing urtication ugli ulceration urticaria vein venipuncture verruca visibility vesiculation vesication vitiligo vellum vernix webfoot waterskin whip wale wattle wineskin welt wanness whitehead wen wind windburn wheal xanthosis xanthoma xeroderma xerodermia zit

Word: instrument

ablative anterior alto amount action act attachment assignment agency adjustment air acceleration activity analysis accompaniment arrangement accelerator accelerometer accordion altazimuth airplane altimeter actinometer aircraft alidade adapter alidad area aperture analyzer analyser arm autofocus atmometer artillery audiometer automobile auriscope auroscope asdic accord atmosphere accuracy altitude angle agent arranger arc azimuth bottom breathing bone bronze bent bowing base baritone bass boot bank burning bow breath browse building bar banjo bagpipe ballistocardiograph band balalaika beam barometer bassoon bend bathymeter bathometer bell board blade binoculars bore body bones bolometer bronchoscope brass burette buret bugle bulb blood brain bean bellow bamboo cut color content current cross criminal clarion chemical cause comparison change completion circle circuit contraction combination contract chair cautery cryocautery charge capacity control class calculation composition contrivance calibration claim case copper chamber cardiograph castanets clappers caliper calliper cangue calliope cabinet console calorimeter camera cauterant ceiling celesta cello chronoscope clinometer cittern cithern cither citole clarinet choir clavier clavichord chime chordophone colorimeter coil column compass component cornet computer contrabassoon contrafagotto concertina cryoscope craniometer curette curet cymbal cloth cymograph curvature cavity cornea corpuscle chord cloud consort craftsman cocuswood cocoswood centrifugation centimeter century derivative drumhead differential drawing device deflection deflexion diffusion division dance dip display direction drum dose diamond declinometer dermatome dashboard densimeter densitometer dilator dilater disk dial dialyzer divider document dosemeter dosimeter dynamometer dulcimer drumstick density distance duration depth datum declination debt dancer entity event execution exposure electrocution elevation examination end echo evaporometer electrocardiograph extractor ergometer electroencephalograph edge electrodynamometer electromyograph esophagoscope endoscope eye engine eyepiece ensemble electroscope euphonium eudiometer earth embouchure energy eyeglasses effect esophagus etude earthquake evaporation fluid force fingering falsification foot flow frame figure function forgery field float fascia firearm fiber film fiberscope finger fingerboard flugelhorn fluegelhorn flute flagpole fret fiddle fluctuation floater flourish fanfare folk family freezing frequency group gash government gas gittern gong glass ground garrote garotte gallows gauge gage garrotte gourd graduate glockenspiel gravimeter grid guillotine guitar goad glasses graticule gravity graph granadilla gamelan glaucoma head hearing hammer hand horn hair hoop housing heliometer hematocrit haematocrit hemostat haemostat heart heat harp harmonica harmonium handle hodoscope hole horsehair hydrometer hygrometer hautboy hautbois horologe hardness height horizon humidity inside intent interior irregular iron interaction implementation instrumentation sentence manner place verb right interest property person voice set performance music piece particle rocket intensity radiation movable part theodolite telescope plane table line sight object musician style measure rate water sensitivity light measuring pulse term time range member torture leg strand wood skimming neck wind period recoil ventricle string pressure percussion use violin pair thumb rhythm tube orchestra section length mean mouthpiece tap tone tissue substance reason removal wart standard point plural punishment steam whistle keyboard radio television quantity input operator system steel plate measurement interval surveyor order inclination soundbox wire key piano sound lute reagent reed melting skull size scoop paper stylus record tree value security pen pointer movement position slice skin panel opening organ indicator solution membrane segment radioactivity power law voltage wave muscle surgery machine warfare ram lense tenor tuba volume reaction player vision solo note sequence placement optic strip woodwind rod red metal tune strangulation thickness rain variation pole state motion network xylophone kettledrum marimba spiral plasma separation star view vessel row physic trace path score skill relationship process role production intonation inclinometer integrator interference pattern interferometer meter user indication idea philosophy success instrumentalism profession instrumentalist kicksorter tortoiseshell kazoo irregularity surface keratoscope tension kettle zither koto tool weapon knife kymograph keyboardist shoulder strap instance vibrato pitch quality keytar limb lagerphone larynx laryngoscope ship speed log luthier mute mark singer reproduction pomp microscope procedure structure microsurgery membranophone magnetograph image magnifier pebble maraca scale resonator plectrum mandolin mallet microtome reading needle velocity nephoscope ocular incision operation orchestration oesophagoscope pipe oboe octant ocarina otoscope retina ophthalmoscope oxygen oximeter melody playing will requirement probate ivory plastic pick investigation injury probe wound planimeter source photometer paddle panpipe pianoforte number step pedometer regulator peg laser photocoagulation photocoagulator pelvimeter periscope post wrist pillory

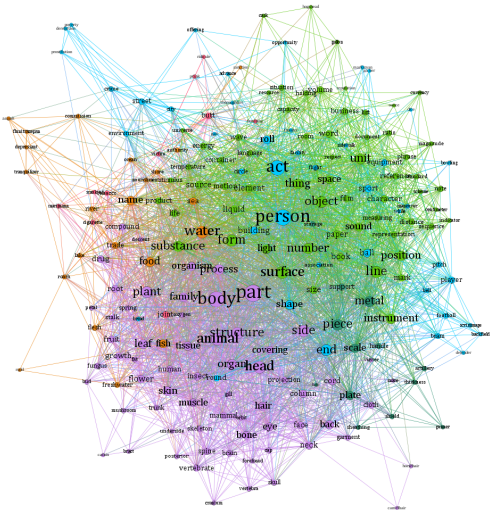
plectron piezometer plethysmograph picture plotter planet planetarium liquid pipet pipette respiration perspiration polygraph potentiometer prod screen projector protractor lyre psaltery precipitation pluviometer polyphony percussionist quadrant victim rack stop timbre register recorder index refractometer rangefinder microwave radar radiolocation reticle reticule retractor rheometer mirror rhinoscope reset kind stopping implement switch slash stab standardization standardisation syrinx sonometer sousaphone saxhorn trombone sackbut samisen shamisen material sclerometer seismograph sextant sensitometer sector tip silverpoint sitar sights slipstick polyp tumor snare spherometer stake recording rapidity spirograph lung spirometer transmission reflection sonograph sonar spectacles specs vibration soundboard spectroscope speculum passage timber offender stocks stroboscope strobe stethoscope pilot propeller synchroscope synchronoscope synchronizer synchroniser variety synthesizer syringe synthesiser semitone strings thrust tripod transit tuning temperament trumpet trump tintometer tympan tympanum timpani tympani inch tape psychologist tachistoscope rotation tachometer tach tapeline information observer telemeter timepiece timekeeper theremin thumbscrew tensiometer moisture soil temperature thermograph thermometer trepan trephine triangle shape tonometer tracing tracer slide tucket run toccata tootle transcriber udometer service voluntary voicing valve violoncello variometer viscosity viscometer viscosimeter vibraphone vibraharp vibes viol viola lash whip wheel watt wattmeter zithern

C.3 Depth Search task

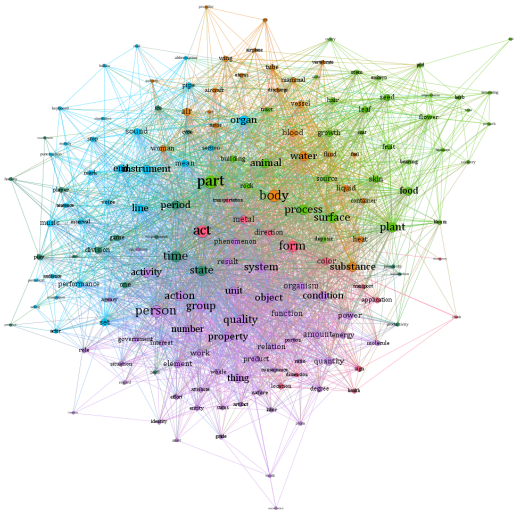
[**Below we present the four sub-networks utilised in *Depth Search*. For details on the implementations and subject behaviour please contact the authors.**]

Figure C.2: Overview of all sub-networks

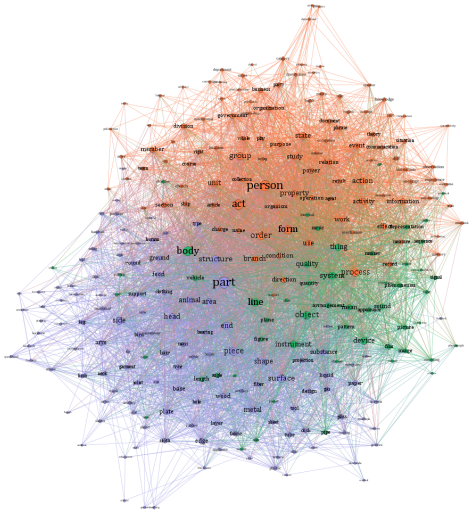
seed: rudd



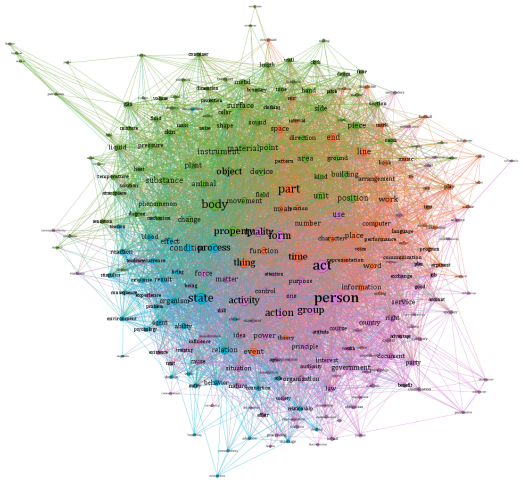
seed: dhal



seed: crinoid



seed: syringe



C.4 Unusual uses task

[**Due to a lack of space, we cannot list all submitted Uses in this section. Please contact the author for the data on Uses submitted.**]

C.5 Tower of London

[**The problems utilised correspond to these network positions (32 - 61), (25 - 54), (34 - 55), (22 - 31)**]

Figure C.3: Problem space of the Tower of London Task by Fimbel et al. (2009).