UNIVERSITY OF TWENTE.

Creative Word Associations:
The Relationship between Creative Idea Generation and Evaluation

Johann Richard Mayer

1st Supervisor: Karolina Rataj, PhD

2nd Supervisor: Prof. Dr. Frank der Velde

Bachelor thesis

Department of Human Factors and Engineering Psychology

Faculty of Behavioural, Management and Social Sciences

July 1, 2022

Abstract

The process of generating a creative word association, one that is both novel and understandable, is assumed to consist of two stages: that of idea generation and that of idea evaluation. The specific underlying cognitive processes in these phases are of special interest in the current research. A recent EEG study by Rataj et al. (2018a) employed an alternate use evaluation task and found effects in the upper alpha band, the N400 amplitudes, and sustained negativity amplitudes related to the evaluation of creative uses and anomalous word pairs. These findings point to the involvement of conceptual expansion, semantic processing, and meaning reintegration in the evaluation phase. Participants of the ERP experiment also completed a word association task, in which they generated creative word associations. In the current study, the word associations were evaluated for novelty and understandability by a new group of participants. The relationship between these ratings and the EEG measures was explored to investigate how the ability to generate creative word associations predicts involvement of processes like conceptual expansion and meaning reintegration during creative idea evaluation. Additionally, it was investigated how mood of participants influences their evaluation of the creativity of others. Correlation and linear regression tests were performed between the novelty and understandability ratings and the obtained EEG measures. No correlations were found between novelty or understandability ratings and measures of alpha desynchronization or N400 amplitudes. Also, no correlation was found between mood of participants and their ratings of creative ability in others. However, a relationship between word association ratings and sustained negativity was found. High understandability of word associations was found to predict less sustained negativity during the evaluation of creative uses and unrelated word pairs. High novelty predicted more sustained negativity during the evaluation phase of creative uses and unrelated word pairs. These findings indicate that creative ability predicts more sustained negativity during the evaluation of creative ideas.

Contents

1. Introduction

A crucial skill for the survival of humankind is the ability to adapt to changes in the environment by generating solutions appropriate to the new circumstances. This skill has not only helped people to survive but also fosters personal growth and plays a major role in professional success (Mastria et al., 2019). Exactly this - generating novel ideas which are appropriate to the situation - is what is widely accepted to be the definition of creativity (Runco & Jaeger, 2012). In the 1930s, the first documented studies about creativity were performed (Guilford, 1967) and to this day, interest in the topic remains. However, creativity involves some of the most complex cognitive functions (Chrysikou, 2019; Guilford, 1967) and little is known about its underlying processes.

In the following, the emergence of theories about the process of generating creative ideas as proposed by some models of creativity will be outlined and followed by an introduction of the two-fold model of creativity by Kleinmintz et al. (2019), expanding the approach of previous models. Subsequently, research concerning the processes involved in creative idea generation and evaluation will be reviewed, including a study by Rataj et al. (2018a), which is going to be the starting point for the research described in this thesis. Furthermore, individual differences influencing the creative process will be highlighted.

1.1 Generating Creative Ideas

Two concepts that were initially found to be important in the process of generating creative ideas were convergent thinking and, more importantly, divergent thinking (Guilford, 1956). While convergent thinking is analytical and more focused on the refinement of one single idea, divergent thinking is associative and focuses on the generation of multiple solutions to a creative problem (Cropley, 2006; Guilford, 1984). Everyday practices like brainstorming or free writing are some examples of divergent thinking which are often used as tools to start the process of creative thinking (Kalargiros & Manning, 2015; Ni et al., 2014). In research, some common measures of creativity are divergent thinking tasks like the Alternate Uses Task (AUT) or Word Association Tasks (WAT) (Sowden et al., 2014). In the former, participants are presented with a word describing a common object and are then asked to come up with an original or creative use for the given object (Vartanian et al., 2019). In the latter, participants are presented with a stimulus word and then asked to respond with a word they associate with the stimulus (Agdam, 2014). Generating novel creative ideas is often considered to involve divergent thinking. However, to be considered creative an idea

must not only be novel but also appropriate. Some early theories suggest that convergent thinking is applied in a process of idea evaluation to assess whether an idea is appropriate (Guilford, 1956). Most of the recent studies abstain from labelling this process as convergent thinking and instead refer to analytical or inhibitory processes (Sowden et al., 2014). Nevertheless, generation of creative ideas involving a phase of idea generation and a phase of idea evaluation, is represented in a multitude of models up to this day (Campbell, 1960; Guilford, 1956; Howard-Jones, 2002; Kleinmintz et al., 2019; Nijstad et al., 2010).

The two-fold model of creativity by Kleinmintz et al. (2019) is one of such models and will be the interpretation of the creative process which this study follows. It proposes a stage of idea generation in which remote associations in the semantic network of individuals are combined in novel ways. Multiple ideas are generated in a defocused manner enabling more original associations. In this stage, the default-mode network, assumed to underlie idea combination, is active. This network is a cluster of different brain regions mainly located in the frontal, temporal, and parietal cortex (Smallwood et al., 2021) and was found to be involved in future planning, mind wandering, retrieval of memories, semantic integration, and divergent thinking (Kleinmintz et al., 2019). In a second stage, these ideas are evaluated. Here, inhibition was found to play a role in rejecting ideas which are not original, allowing for more remote associations to be formed. However, in opposition to previous models of creativity, Kleinmintz et al. (2019) propose that the evaluation stage does not map directly onto executive control processes as, for example, convergent thinking does. Instead, it is assumed that there are three substages of evaluation which involve different types of processes. First, the valuation process, relating to emotional and motivational processing, in which the appropriateness is evaluated. Second, the monitoring process, in which the novelty is evaluated, and which relates to higher executive control processes, responsible for inhibition. Third, the selection process, in which ideas are selected based on the combined outcome of the two previous processes.

Furthermore, the two-fold model suggests that the two phases of idea generation and idea evaluation interact in a cyclic motion. In the successful production of creative ideas there seems to be a close link between the activity in the default-mode network (DFM) and the executive control network (ECN). The ECN is a network of brain regions which lie mostly in the dorsolateral prefrontal cortex as well as the dorsal anterior cingulate cortex. It is strongly associated with cognitive control processes like inhibition and is responsible for directing attention (Ellamil et al., 2012; Miller & Cohen, 2001. Due to this interplay between idea

generation and evaluation, Kleinmintz et al. (2019) suggest that the stringency of evaluation influences the phase of idea generation. An evaluation that is too stringent can inhibit the generation of original ideas while an evaluation that is too lenient might result in ideas which are not original.

1.2 Neural Activity During Generation of Creative Ideas

To investigate the neural underpinnings of the process of generating creative ideas and this underlying cyclic motion, previous research has often employed electroencephalography or transcranial direct (or alternating) current stimulation. Electroencephalography (EEG) is used to measure the bioelectrical signal evoked by synchronous activity of brain cells through nodes placed on the scalp picking up activity of electric fields. Oscillations of different frequencies are distinguished and associated with different mental states of the subject. Alpha band oscillations are measured from 8-13Hz and are of particular interest in research of the creative process as studies suggest a correlation between alpha band activity and divergent thinking (Benedek et al., 2011). Another point of interest regarding creativity research is the N400. The N400 is an event-related potential (ERP) observed in response to stimuli such as words or pictures. It is a negative-going wave which peaks at 400ms after the presentation of a stimulus containing semantic information and its amplitude depends on semantic difficulty as well as semantic novelty of the stimulus (Rataj et al., 2018b). This is relevant in the process of conceptual expansion, which is a key part to generating creative ideas. Conceptual expansion is the ability to expand the understanding of a concept by adding new associations (Ward, 1994). Following the N400 sustained negativity has been observed in ERP studies on creativity. It is a negative-going wave observed between 500 and 900ms after the stimulus is presented. It is interpreted as indicating processes of meaning integration as well as, when measured over anterior sites, demands on the semantic working memory (Steinhauer et al., 2010).

Another commonly used method to investigate the cognitive processes active in creative idea generation is transcranial direct current stimulation(tDCS). This is a method of non-invasive brain stimulation sending small electric currents of 1 - 2 mA through the skull, to alter cortical excitability. During a creative task, cortical areas can be either excited through anodal stimulation or inhibited through cathodal stimulation, to assess the involvement of the targeted areas in the creative process. The frontotemporal areas have been of particular interest in previous studies on creativity as these are hypothesized to regulate

cognitive control during divergent thinking tasks (Weinberger, 2017). A similar method is that of transcranial alternating current stimulation (tACS). While tACS does not change cortical excitability like tDCS does, it can be used to entrain neurons to fire at a desired frequency.

Two recent studies by Chrysikou et al. (2021) and Kenett et al. (2021) have used tDCS to investigate the involvement of inhibitory processes in creative thinking. They have found that changes in stimulation of the left prefrontal cortex (LPC) during a divergent thinking task correlated with quality and quantity of creative responses. While the study of Chrysikou et al. (2021) showed that inhibitory stimulation of the LPC resulted in a more fluent phase of idea generation, Kenett et al. (2021) found that excitatory stimulation of the LPC increased appropriateness and decreased novelty of responses. This points to the importance of inhibitory processes during divergent thinking, further supporting the idea of the cyclic motion between idea generation and idea evaluation. A similar method is used in a study by Lustenberger et al. (2014). Employing tACS allowed Lustenberger et al. (2014) to target alpha oscillations and enhance alpha power in the frontal cortex. Participant's creativity was measured under the exposure to tACS, enhancing alpha power. Their results showed an increase in creative ability, pointing to the involvement of alpha band activity in the creative process.

Similarly, studies employing EEG have found an increase in alpha band oscillations during the creative process (Benedek et al., 2011; Fink & Benedek, 2014). Such an increase in alpha band activity is commonly related to the involvement of top-down processes found to be relevant for the suppression of irrelevant information, directing attention, and high-level semantic processing (Camarda et al., 2018; Luft et al., 2018; Rataj et al., 2018a). The importance of alpha band activity and inhibition in the process of creating semantic associations is highlighted in research reported by Luft et al. (2018). In their study, four experiments were performed in which participants completed either a convergent or a divergent thinking task while either having their EEG measured or being exposed to tACS. Luft et al. (2018) have found that participants exposed to tACS generated more creative ideas, replicating the findings of Lustenberger et al. (2014). Furthermore, the EEG showed that participants not exposed to tACS showed higher alpha power during the generation of more creative ideas. The results of these experiments suggest that an increase in alpha oscillations during idea generation may support an individual in inhibiting semantic associations which are obvious, leading to more creative ideas.

Instead of focussing on the process of creative ideation, a recent study by Rataj et al. (2018a), focussed on the role of alpha band oscillations, N400 amplitudes, and sustained negativity during the process of creative idea evaluation. In this study, participants performed an alternate use evaluation task (AUeT) in which they were presented with word pairs representing common uses (e.g., flowers as a crown), creative uses (e.g., nest as a crown) and unrelated word pairs (e.g., violin as a crown). Participants evaluated whether the presented use is common or uncommon and how usable it is. Participants' EEG was measured during these tasks and alpha band activity, and amplitudes in the N400 and sustained negativity were analysed. An increase in the activity of the upper alpha band was found during evaluation of creative uses and unrelated word pairs, indicating increased demands in accessing and integrating semantic information. Furthermore, an increase in activity of the lower alpha band during the evaluation of creative uses and unrelated word pairs was found and hypothesized to indicate higher attentional demands. This falls in line with various findings relating an increase in alpha power to internally directed attention (Benedek et al., 2018). Also, an increase in N400 amplitudes was found during evaluation of creative uses and unrelated word pairs which is hypothesized to indicate an increased demand in the retrieval of semantic information for conceptual expansion.

Various studies investigated the N400 through means of metaphor comprehension (Rataj et al., 2018b; Goldstein et al, 2012). In these studies, participants were presented with literal metaphoric utterances (e.g., a new theory) and novel metaphoric utterances (e.g., a moldy theory) and had to determine their meaningfulness. These studies have found that evaluation of novel metaphoric utterances correlates with larger N400 amplitudes. This might be attributed to the need to create new meaningful associations in the comprehension of novel metaphoric utterances, while literal metaphoric utterances require only the retrieval of meaning. Creating these new associations appears to be achieved through gathering and recombining semantic information in a process of conceptual expansion (Rataj et al., 2018a).

However, not only increased N400 amplitudes are commonly associated with evaluation of stimuli containing new semantic information, a relationship with sustained negativity has also been found in several studies (Goldstein, 2012; Rataj et al., 2018b). An EEG study by Rutter et al. (2012) investigated the underlying cognitive processes during creative idea evaluation by presenting participants with novel, literal, and nonsensical metaphoric phrases with the task to evaluate their unusualness as well as appropriateness.

EEG data showed higher N400 amplitudes for phrases that were evaluated as unusual and inappropriate as well as those evaluated as unusual and appropriate, compared to amplitudes measured during evaluation of literal phrases. This indicates higher demands for semantic information retrieval and conceptual expansion. The same effect was also recorded for the sustained negativity, only that negativity of amplitudes recorded for novel metaphoric phrases decreased over time while those for nonsensical phrases remained negative. In a study by Jiang et al. (2009), sustained negativity without a preceding increase in N400 amplitudes was found when participants were presented with sentences containing a misused universal quantifier ("He threw away all that apple"). These findings support the assumption that sustained negativity is associated with handling nonsensical semantic input and meaning reintegration. It is hypothesized that the sustained negativity represents a process of reinterpretation when meaning could not be found after the initial attempt of semantic integration (Jiang et al., 2019).

A study by Kenett et al. (2018) highlights the importance of a flexible semantic memory structure for the ability to comprehend novel metaphors. In this study a group of high creative individuals and a group of low creative individuals were asked to judge semantic relatedness of novel, conventional, literal, and unrelated metaphoric word pairs. They have found that high creative individuals were quicker in detecting both novel as well as conventional metaphoric word pairs and more accurate in detecting the former. While quickness can be attributed to executive functions (Benedek et al., 2014), accuracy might best be accounted for by individual differences in semantic memory structure (Kenett et al., 2018; Kenett et al., 2014). While the semantic memory structure of low creative individuals appears to be more rigid, that of high creative individuals is more spread-out (Kenett et al., 2018).

1.3 Motivational and Emotional Processes

Besides the involvement of inhibitory processes, motivational and emotional processes were highlighted in the two-fold model and must not be disregarded (Kleinmintz et al., 2019). Individual differences in these domains have also been investigated and were found to influence the ability to generate and evaluate creative ideas. In a study by Benedek et al. (2016), which asked participants to evaluate 72 ideas as common, inappropriate, or creative, it was found that creativity as well as creative achievement correlate with evaluation skills. Participants in this study underestimated the creativity of others, supporting a body of research suggesting that individuals underestimate the creativity in ideas of others compared

to their own ideas (Grohman et al., 2006; Silvia, 2008). Furthermore, Benedek et al. (2016) suggest that this underestimation might be due to a lack of understanding. However, openness to experience, intelligence, and language competence were found to predict less underestimation. Meanwhile, in a study by Mastria et al. (2019), participants were asked to evaluate the ideas of others while being in different emotional states. Participants who were in a more positive mood had the tendency to evaluate ideas as more creative, while those in a negative mood evaluated them as less creative. These findings potentially further support the two-fold models claim of the involvement of emotional processes in the evaluation phase.

1.4 The Current Study

Previous EEG studies found an increase in alpha power during the generation of creative ideas (Luft et al., 2018, Fink & Benedek, 2014) and studies employing tACS or tDCS could even increase the ability to generate creative ideas through increasing alpha band activity (Luft et al., 2018; Lustenberger et al., 2014, Kennet et al., 2021). This indicates semantic information retrieval and demands for internally directed attention during generation of creative ideas. Rataj et al. (2018a) were the first to investigate upper alpha ERD during an alternate use evaluation task and found that evaluation of creative uses and anomalous word pairs evokes more alpha ERD than evaluation of common uses. Furthermore, Rataj et al. (2018a) found more N400 amplitudes and sustained negativity during the evaluation of creative uses and anomalous word pairs. This can also be found in other studies investigating N400 and sustained negativity during the evaluation of creative ideas and points to the involvement of processes like conceptual expansion and meaning reintegration (Rataj et al., 2018b, Rutter et al., 2012). As the two-fold model (Kleinmintz et al., 2019) proposes evaluation to be a crucial part of generation of creative ideas, it can be hypothesized that there is a relationship between these EEG measures of alpha power, N400, and sustained negativity during the evaluation of creative ideas and the ability to generate creative ideas. However, this relationship is yet to be fully discovered and explained.

This study was the first to investigate the relationship between the ability to generate creative ideas and alpha desynchronization, N400 and sustained negativity during evaluation of creative ideas. For this, EEG measures and word associations obtained from the previously mentioned study by Rataj et al. (2018a) were used. This EEG study by Rataj et al. (2018a) measured alpha desynchronization, N400 amplitudes and sustained negativity during an AUeT. In this task, participants were presented with different types of word pairs

representing common uses, creative uses, as well as unrelated word pairs. Then, these participants evaluated whether this use was common or uncommon and, on a scale from one to four, indicate how usable it was. Afterwards, without having their EEG measured, these same participants completed a word association task. In the word association task participants were presented with a stimulus word and gave a creative association to this stimulus word. Also, they gave an explanation to why they have associated these two words.

In this current study, these word associations were used to create a rating-scale study to evaluate the creative ability of participants of the word association task. A new group of participants was recruited for the rating-scale study. To avoid confusion between the two different participants groups, participants from the EEG study by Rataj et al. (2018a) will be referred to as WA (word association task) participants and participants from this current study will be referred to as RS (rating-scale study) participants. In this current rating-scale study, RS participants were presented with word associations and their respective explanations created by WA participants of the study reported by Rataj et al. (2018a) and evaluated them on two scales measuring novelty and understandability. As previous studies indicate that mood influences the evaluation of creativity of others, the PANAS mood scale (Peeters et al., 1996) was also added and filled out by these participants.

Considering that an increase in alpha desynchronization was found during the generation and evaluation of creative ideas and the cyclic motion of these phases as proposed by the two-fold model, the first research question arises: Does the ability to generate creative word associations predict smaller alpha desynchronization during the evaluation of creativity?

As an increase in N400 amplitudes was found in the evaluation of creative and anomalous word pairs but the relationship between this measure and the ability to generate creative ideas is yet unclear, the second research question will be: What is the relationship between the ability to generate creative word associations and N400 amplitudes during the evaluation of creativity?

Similarly, an increase in sustained negativity was found in the evaluation of creative and anomalous word pairs but the relationship between this measure and the ability to generate creative ideas is also still unclear, leading to the third research question: What is the relationship between the ability to generate creative word associations and sustained negativity during the evaluation of creativity?

Furthermore, in past studies mood of participants was found to influence their evaluation of creativity of others, hence the fourth research question will be: Does mood of rating-scale study participants influence their evaluation of creative word associations?

To answer these questions, the following hypotheses were tested:

H_1: Alpha desynchronization recorded in WA participants during the AUeT is predicted by the creativity of their word associations as assessed through the RS participants.

H_2: A significant correlation will be found between N400 amplitudes recorded in WA participants during the AUeT and the creativity of their word associations as assessed through the RS participants.

H_3: A significant correlation will be found between sustained negativity recorded in WA participants during the AUeT and the creativity of their word associations as assessed through the RS participants.

H_4: The PANAS mood score of RS participants predicts how novel and creative these participants rate the word associations presented to them.

2. Method

2.1 Design

The study was set up as a cross-sectional survey following a questionnaire survey design. Correlation and simple linear regression tests were performed to investigate the relationship between EEG measures of WA participants during an AUeT and the novelty as well as understandability of their word associations as assessed RS participants. Furthermore, correlation tests were performed to analyse the relationship between mood of the RS participants and their evaluations of the novelty and understandability of the word associations.

2.2 Materials

2.2.1 EEG Data and Word Associations

EEG data were obtained from a previous study by Rataj et al. (2018a). In this study, the N400 amplitudes, sustained negativity, as well as oscillations in the upper alpha band were obtained from 22 participants during an AUeT. The upper alpha band measures used in this study are those recorded over parieto-occipital sites four and eight (PO4 and PO8). Measures of N400 amplitudes and sustained negativity are those acquired over central and frontal sites. All measures were obtained during each of the three conditions: evaluation of creative uses of objects, evaluation of common uses of objects, and evaluation of unrelated word pairs.

Furthermore, a total of 400 word-associations were obtained from the study by Rataj et al. (2018a). After completing the AUeT, the 22 WA participants were instructed to give a creative association to 21 stimulus words each, accompanied by an explanation for each association. Before their first association, each WA participant was given a word to practice on to ensure that the task was clearly understood and only began when the researchers knew that it was clear to the WA participants what a creative association was.

2.2.2 Rating-scale Study

The word associations alongside their respective explanations were split into ten sets. Each set consisted of two blocks containing between 17 and 21 associations from various WA participants, separated by a short break. In each set, every stimulus word was shown twice, once in each block. This was done so that RS participants would not see multiple associations to one stimulus, preventing their judgement to be biased from seeing other

associations to the same stimulus. In some cases, WA participants did not create a word association for the stimulus word, therefore in some blocks, some stimulus words did not appear.

Furthermore, twelve filler items (Appendix A) were added to each set of which six were very original but not very understandable and six were not very original but very understandable. Three fillers of each type were added to block one and another three of each to block two. This was done so that the far ends of each scale would be established in each block. This way, RS participants could get an idea of what a very novel or very understandable association would look like and therefore evaluated other word associations with vivid examples of both ends of the scale to refer to. This resulted in ten sets containing two blocks of 23 to 27 associations including fillers. Then, for counterbalancing, a copy of each set was created and the order of the blocks in the copied set was reversed, resulting in a total of 20 sets. While, for example, Set A_1 would show block one first and block two after the break, Set A_2 would show block two first and block one after the break. This counterbalancing of sets within blocks was done to prevent an order effect or fatigue from influencing the judgement of RS participants as all items were equally distributed across both halves of the survey. To further prevent a serial order effect, the order of items in each block was randomized.

2.2.3 Mood Measure

To investigate the influence of mood on the ability to evaluate the associations, a Dutch version of the PANAS mood scale (Peeters et al., 1996) was included in the beginning of the study (Appendix B). The RS participants were asked how much they agreed with ten statements about their current mood on a five-point Likert scale. Five items targeted positive mood while the other five were focused on negative mood. The PANAS was chosen as it has proven to be an effective tool in previous studies about the effect of mood on creative idea evaluation (Mastria, 2019). In the present study, the Dutch adaption of the PANAS by Peeters et al. (1996) was used which was found to be the most valid and reliable translation of the scale (Engelen et al., 2006). Reliability is indicated by a Cronbach's alpha of 0.83 for the five negative items and of 0.79 for the five positive items. Validity of the five positive and the five negative items is shown through a factor analysis resulting in an explained variance of 38.7%.

2.3 Participants of the Rating-Scale Study

The participants of the rating-scale study were 131 Dutch native-speakers between 18 and 60 years of age (Mean = 26.01). RS Participants were excluded from the study if they were below 18 years of age, native speakers of languages other than Dutch, had a reading or language disorder, did not give informed consent, did not complete the study, or previously participated in the study "Evaluation of creative word associations". Based on these criteria, 66 participants had to be excluded, leaving a sample size of 57, where each of the ten sets had between five and six participants, except for set D, which was completed by eight participants. The mean age of these 57 participants was 24.96. When asked for their gender, 25 indicated their gender as male, 24 as female, and one as "other". Of these participants, the educational level of one (2%) was a doctorate degree, eight (16%) had a master's degree, 26 (52%) had a bachelor's degree, and 15(30%) VMBO, HAVO, VWO or MBO. The intended sample size was 160, with sixteen participants per set. Before recruitment of participants, ethical approval of the Ethics Committee at the Faculty of Behavioral Sciences at the University of Twente was obtained.

2.4 Procedure

RS Participants were led to the study through a Qualtrics link distributed via social media alongside an invitation message or joined through SONA, a participant recruitment platform of the University of Twente. Questions regarding the exclusion criteria for this study were asked and informed consent as well as demographic information was obtained. Furthermore, RS participants were informed about the duration of the study, confidential handling of their data and the right to withdraw at any time. Also, it was emphasized that there are no correct answers, as only the RS participants opinion matters. Next, the RS participants mood was inquired with 20 questions of the PANAS mood questionnaire. Then, RS participants were given instructions on what to do. They were introduced to three example associations and the two seven-point Likert scales to rate originality and understandability (Appendix C). The scales were explained, and RS participants were asked to rate the three examples on the originality scale ranging from "het is zeer gewoon" (it is very common) to "het is zeer vernieuwend" (it is very novel). The understandability scale ranged from "het is zeer onbegrijpelijk" (it is very incomprehensible) to "het is zeer begrijpelijk" (it is very understandable). An example word association with the two scales as presented to the RS participants can be seen in Figure 1. RS participants were then assigned

randomly to one of the 20 sets. In the following, RS participants were asked to evaluate the first block of items, containing about 25 associations, one after the other in random order. Afterwards, the second block was presented. After completion of the two blocks RS participants who joined through SONA were asked for their SONA ID to obtain their credits and the survey ended. Completion of the survey was expected to take about 25 minutes and RS participants were informed that participation is voluntary, they may stop at any time and their data would be handled anonymously.

Figure 1

Example of one word association alongside its explanation and the two scales as shown to RS participants

2.5 Data Analysis

For analysis of the data, different datasets had to be created as the analysis of mood in relation to creativity ratings required data to be set up differently than the analysis of EEG data in relation to creativity ratings.

The dataset of the rating-scale study conducted in Qualtrics was downloaded from the platform and RS participants data was deleted based on the previously mentioned exclusion criteria. Furthermore, data of two WA participants were excluded as they created word associations alongside the explanation that the two words have nothing to do with each other (e.g., Mother – Football. Explanation: They have no connection). As the creation of such word association does not evoke semantic reintegration or retrieval of semantic information but is rather random, answers of these WA participants were deemed unusable for the analysis of the relationship between EEG data and creativity ratings.

To investigate the relationship between EEG data and creativity scores of WA participants, a dataset was created in which WA participants were assigned to the left-most column and their respective EEG measures were in the following columns (Appendices D-F). EEG measures were obtained from Karolina Rataj who gathered these in a study that also included the word association task (Rataj et al., 2018a). In the upper alpha band, the strongest effects were observed over parieto-occipital sites four and eight (PO4 and PO8), and for the N400 and sustained negativity amplitudes over frontal and central sites. For each of these, a measure during evaluation of common uses of objects, a measure during evaluation of creative uses of objects, and a measure during evaluation of unrelated word pairs was obtained and added to the dataset with the WA participants. This resulted in a dataset with 20 WA participants in the left column and, for each WA participant, the EEG measures obtained over different sites during evaluation of creative uses, common uses and unrelated word pairs (Appendices D-F). Then, the mean novelty, as well as the mean understandability scores of word associations created by each WA participant were calculated and added for each WA participant. For this, the respective word associations of each WA participant had to be gathered from the different sets of the rating-scale study as the associations of each WA participant were distributed across multiple sets. An Excel sheet was created in which each word association was assigned to the respective WA participant that created the association. A mean novelty and a mean understandability score for each word association by the WA participants was created based on the rating of the RS participants and added to the table. From these scores, a mean novelty score, and a mean understandability score of each WA participant were created and added to the dataset with their EEG measures. In SPSS, this dataset was analysed by first performing tests of normality, correlation analyses, and tests to check whether the necessary assumptions are met. Then, simple linear regressions were performed for significant correlations found for EEG data with novelty and understandability scores.

Furthermore, a median split was performed to investigate individual differences. For this, the sample of WA participants was split at the median of the novelty scores to create a low-novelty and a high-novelty group. Another split was made at the median of the understandability scores to create a low-understandability and a high-understandability group. Each of the resulting four groups was put in a separate dataset and correlation analyses were performed in each dataset. Simple linear regression was performed for EEG measures that significantly correlated with only novelty or only understandability. simple linear regression

was performed for EEG measures that significantly correlated with both novelty and understandability.

To analyse the relationship between mood and creativity ratings, multiple datasets of a different structure had to be created (Appendix G). The creativity ratings for this analysis do not refer to the mean understandability and novelty of each WA participant, but rather to the novelty and understandability mean scores of each RS participants for the word associations they evaluated. As the dataset provided by Qualtrics showed RS participants in the rows, this dataset could be used to compute the mean novelty and mean understandability rating given by each RS participant. As RS participants were assigned to different sets of word associations and each set contained different items, means could not be compared across sets but only within each set. Therefore, a dataset for each set was created with only the RS participants that evaluated items of this set and their respective understandability and novelty means. This resulted in ten datasets, each containing between five and eight RS participants. In SPSS, each dataset was analysed by first performing tests of normality, correlation analyses and tests to check whether the necessary assumptions are met.

3. Results

3.1 Creative Ability and EEG Measures

3.1.1 Tests for normal distribution of data

To test for normal distribution of data in all variables a Shapiro-Wilk test was performed. Alpha desynchronization during evaluation of common uses of objects measured over the parieto-occipital right electrode sites four (W (20) = .93, p = .13) and eight (W (20) = .95, p = .36) were normally distributed. Data measured over right electrode site four during evaluation of creative uses of objects (W (20) = .77, p < .05) or unrelated word pairs (W (20) = .90, p < .05) were not normally distributed. Data measured over right electrode site eight during evaluation of creative uses of objects (W(20) = .90, p < .05) or unrelated word pairs (W (20) = .90, p < .05) were also found to be not normally distributed.

N400 amplitudes, as well as sustained negativity, measured over frontal and central sites were found to be normally distributed for measures during evaluation of common uses of objects, creative uses of objects, and unrelated word pairs. Mean ratings of novelty (W (20) = .92, p = .09) as well as understandability (W (20) = .93, p = .14) were also both found to be normally distributed.

3.1.2 Correlation tests

For the normally distributed variables, Pearson correlations were computed, while Spearman's rank-order correlations were computed for variables not distributed normally (Table 1). No significant correlations were found between any of the measures of alpha desynchronization and mean novelty and understandability ratings (ps > .05). Also, no significant correlations were found between any of the measures of N400 amplitudes with mean novelty and understandability ratings (ps > .05).

Regarding sustained negativity, a significant negative correlation was found between mean novelty ratings and amplitudes measured over central sites during evaluation of creative (r (18) = -.67, p < .01) and unrelated (r (18) = -.49, p < .01) word pairs (Figure 2). The higher the novelty ratings, the more negative the sustained negativity amplitudes were. A significant positive correlation was found between understandability ratings and amplitudes measured over central sites during evaluation of creative (r (18) = .60, p < .01) and unrelated (r (18) = .59, p < .01) word pairs (Figure 3). Higher understandability ratings correlated with less negative amplitudes.

Furthermore, a significant negative correlation (Figure 4) was found between the two predictor variables novelty and understandability (r (18) = - .90, p < .01). To ensure linearity and homoscedasticity of these correlations, residual and scatter plots were created and found to indicate linearity and homoscedasticity for all correlations.

Figure 2

Correlation between mean novelty ratings and sustained negativity measured over central sites during evaluation of creative uses and unrelated word pairs

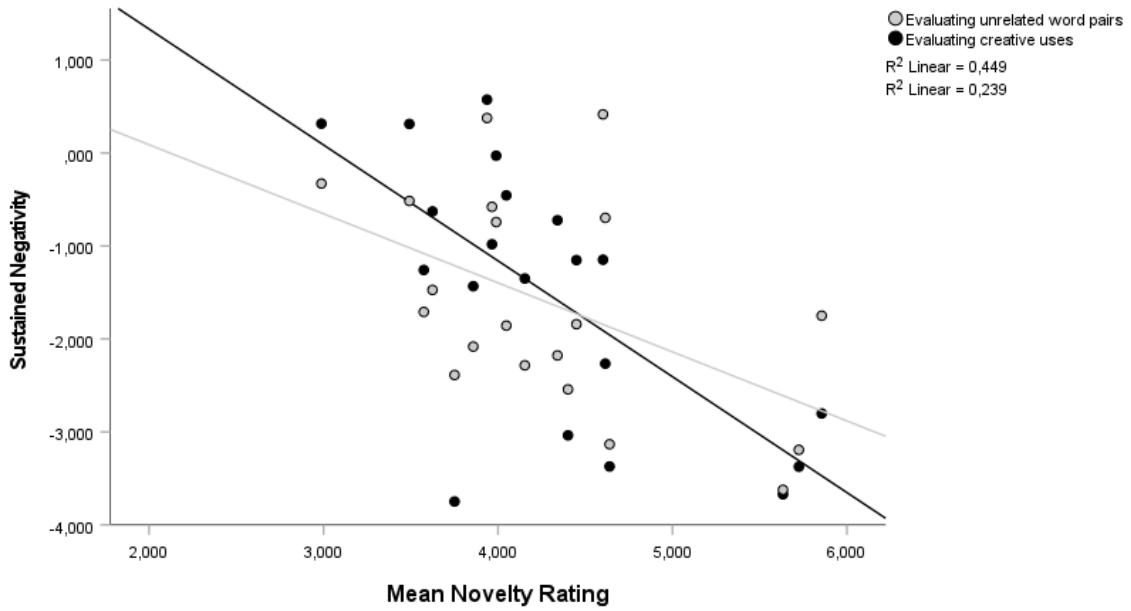


Figure 3

Correlation between mean understandability ratings and sustained negativity measured over central sites during evaluation of creative uses and unrelated word pairs

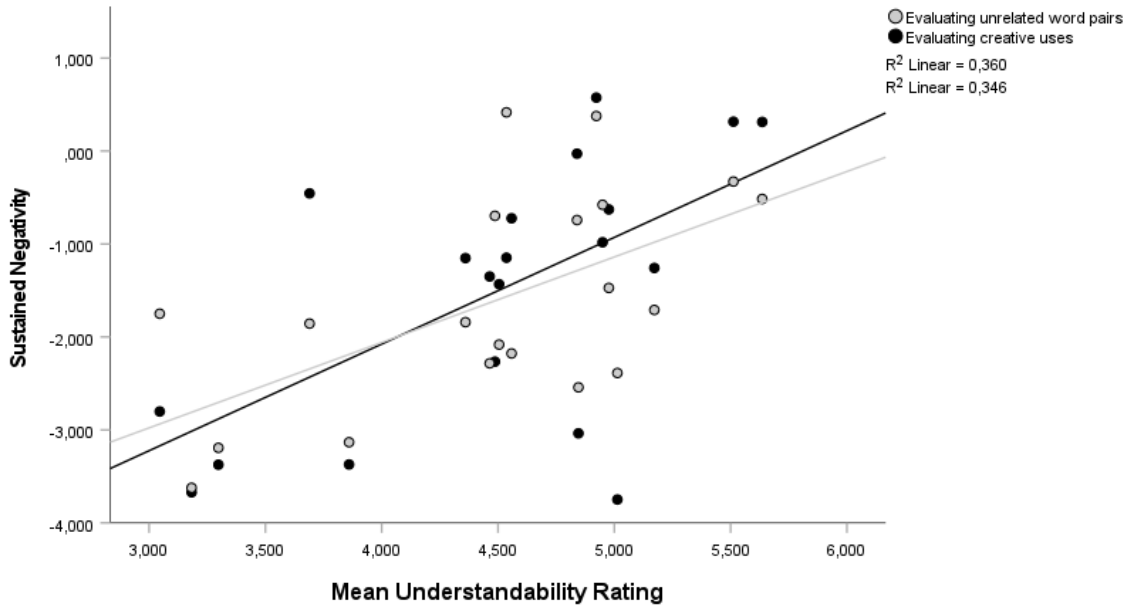


Figure 4

Correlation between mean understandability and mean novelty ratings

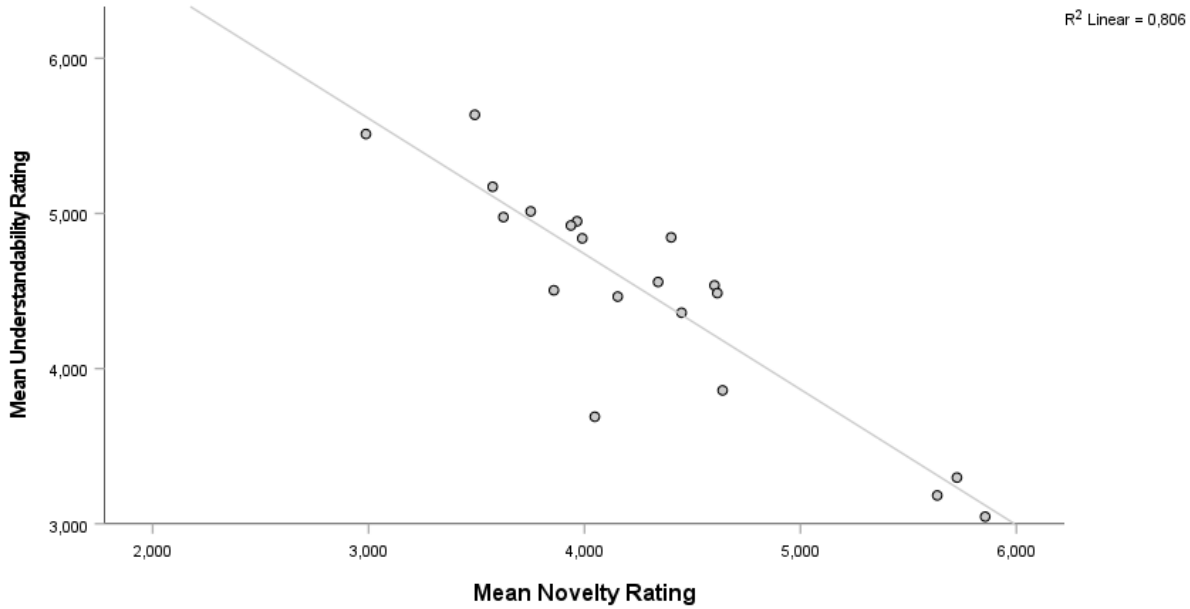


Table 1

Correlation coefficients for the relationship between alpha desynchronization measured over parieto-occipital sites (PO4 and PO8), N400 amplitudes measured over frontal (N400F) and central sites (N400C), sustained negativity measured over frontal (SNF) and central sites (SNC) and mean novelty, as well as mean understandability ratings

	Mean novelty	Mean understandability
PO4_Common uses[a]	-.00	-.02
PO4_Creative uses[b]	-.19	-.06
PO4_Unrelated word pairs[b]	.07	-.18
PO8_Common uses[a]	-.27	.27
PO8_Creative uses[b]	-.27	.21
PO8_ Unrelated word pairs [b]	-.21	.05
N400F_Common uses[a]	.18	-.36
N400F_Creative uses[a]	-.05	-.18
N400F_ Unrelated word pairs [a]	-.04	-.15
N400C_Common uses[a]	.03	-.02
N400C_Creative uses[a]	-.27	.16
N400C_ Unrelated word pairs [a]	-.07	.10
SNF_Common usesm[a]	.35	-.32
SNF_Creative uses[a]	.07	-.09
SNF_ Unrelated word pairs[a]	.06	.74
SNC_Common uses[a]	-.38	.40
SNC_Creative uses[a]	-.67**	.60**
SNC_ Unrelated word pairs [a]	-.49*	.59**

*$p < .05$. **$p < .01$. [a] Pearsons r. [b] Spearman's r

3.1.3 Regression

Two hierarchical multiple linear regression tests were performed with novelty ratings in the first stage and understandability ratings added in the second stage. One was performed for sustained negativity measured over central sites during evaluation of creative uses of objects and one for sustained negativity measured over central sites during evaluation of unrelated word pairs measured over the same sites. As multicollinearity was found between the two predictor variables (VIF = 5.15), simple linear regression models were chosen instead.

A simple linear regression model with sustained negativity measured over central sites during evaluation of creative uses of objects as the outcome variable and novelty as predictor variable was significant ($r^2_{adjusted}$ = .42, F (1,18) = 14.65, p < .01), where novelty ratings were found to significantly predict the outcome variable (ß = -.67).

A simple linear regression model with sustained negativity measured over central sites during evaluation of creative uses of objects as the outcome variable with understandability as predictor variable was significant ($r^2_{adjusted}$ = .32, F (1,18) = 10.12, p < .01). Understandability ratings were found to significantly predict the outcome variable (ß = .60).

A simple linear regression model with sustained negativity measured over central sites during evaluation of unrelated word pairs as the outcome variable and novelty as predictor variable was significant ($r^2_{adjusted}$ = .20, F (1,18) = 5.64, p < .05). Novelty ratings were found to significantly predict the outcome variable (ß = -.49)

A simple linear regression model with sustained negativity measured over central sites as the outcome variable during evaluation of unrelated word pairs and understandability as predictor variable was significant ($r^2_{adjusted}$ = .31, F (1,18) = 9.53, p < .01). Understandability ratings were found to significantly predict the outcome variable (ß = .59)

3.1.4 Median Split

To further explore the data for an influence of individual differences in creative ability on alpha desynchronization and amplitudes in the N400 or sustained negativity, a median split was performed on mean novelty ratings as well as mean understandability ratings.

Results of the Shapiro-Wilk test indicated that understandability was normally distributed in all groups and novelty was normally distributed in all groups except the high novelty group. Pearson correlations were computed between all normally distributed variables and novelty

and understandability ratings. Spearman's rank-order correlations were computed for the variables in the high novelty group with mean novelty and understandability ratings. No significant correlations were found for WA participants in the low-novelty and high-novelty groups (*ps* > .05).

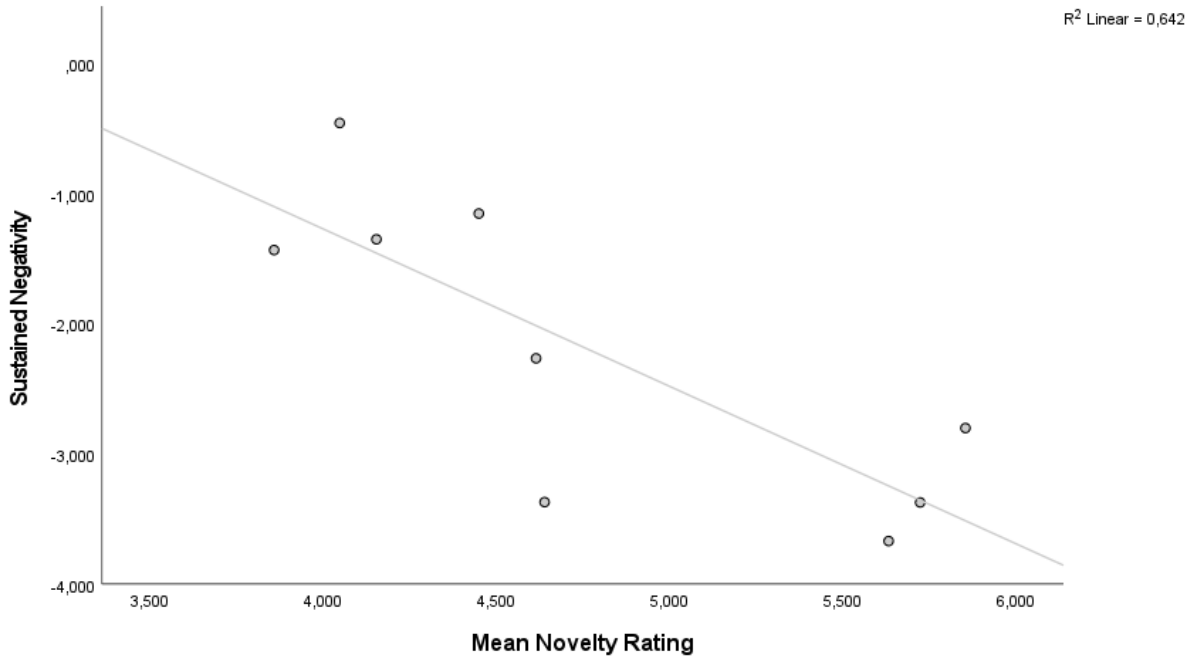
In the group of low-understandability WA participants, a significant negative correlation (Figure 5) was found between novelty ratings and sustained negativity measured over central sites during evaluation of creative uses of objects (r (9) = -.80, p < .01). Higher novelty ratings correlated with more negative amplitudes.

In the group of low-understandability WA participants, a simple linear regression model for sustained negativity measured over central sites during evaluation of creative uses of objects as the outcome variable and novelty ratings as predictor variable was found to be significant ($r^2_{adjusted}$ = .59, F (1,7) = 12.60, p < .01), where novelty ratings were found to significantly predict the outcome variable (ß = -.80).

In the group of high-understandability WA participants, sustained negativity measured over frontal sites during evaluation of common uses of objects was found to have a significant positive correlation with novelty ratings (r (11) = .71, p = .01) and a significant negative correlation with understandability ratings (r (11) = -.74, p < .01) (Figure 6, Figure 7). Also, a significant negative correlation was found between sustained negativity measured over frontal sites during evaluation of creative uses of objects and understandability ratings (r (11) = -.68, p < .05). Residual- and scatterplots indicated linearity and homoscedasticity for all correlations. As multicollinearity in hierarchical multiple linear regressions indicated a possible moderate relationship also in this group (VIF = 4.50), simple linear regressions were performed.

Figure 5

Correlation between mean novelty ratings and sustained negativity measured over frontal sites during evaluation of creative uses in the low understandability group



In the high-understandability group, a simple linear regression model for sustained negativity measured over frontal sites during evaluation of creative uses as outcome variable and understandability ratings as predictor variable was significant ($r^2_{adjusted}$ = .40, F (1,9) = 7.65, p < .05), where understandability ratings were found to significantly predict the outcome variable (ß = -.68).

In the high-understandability group, a simple linear regression model for sustained negativity measured over frontal sites during evaluation of common uses of objects as outcome variable and novelty ratings as predictor variable was significant ($r^2_{adjusted}$ = .45, F (1,9) = 9.13, p = .01). Novelty ratings were found to significantly predict the outcome variable (ß = .71).

In the high-understandability group, a simple linear regression model for sustained negativity measured over frontal sites during evaluation of common uses of objects as outcome variable and understandability ratings as predictor variable was significant ($r^2_{adjusted}$ = .50, F (1,9) = 11.01, p < .01). Understandability ratings were found to significantly predict the outcome variable (ß = -.74).

Figure 6

Correlation between novelty and sustained negativity measured over frontal sites during evaluation of common uses in the high understandability group

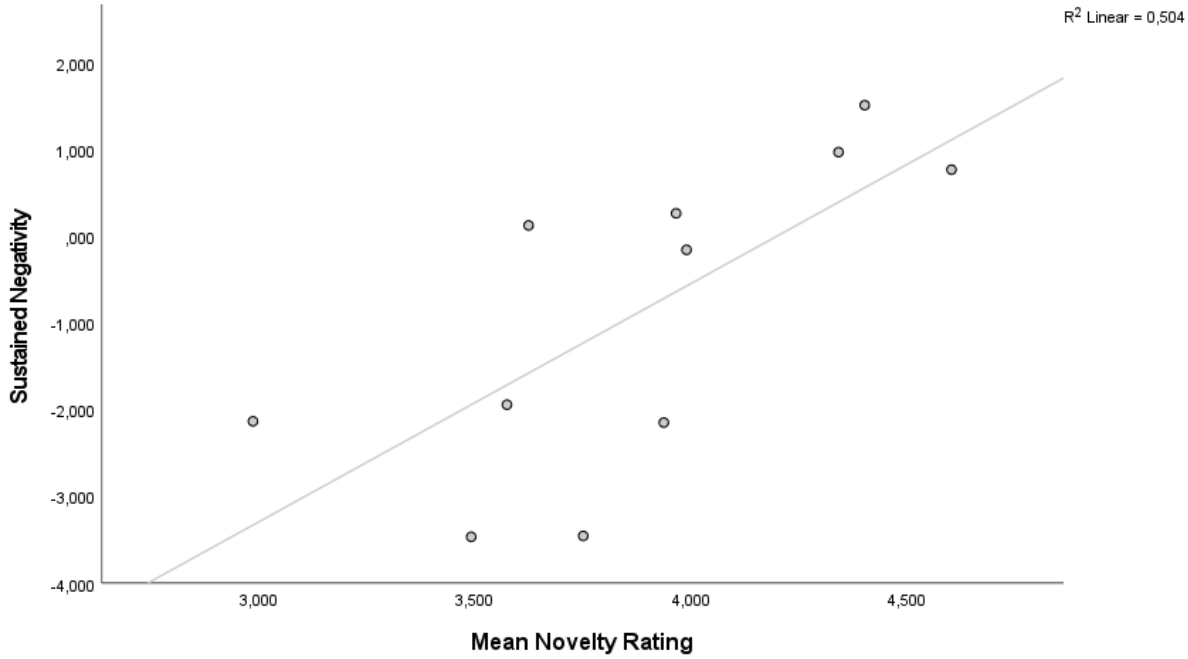
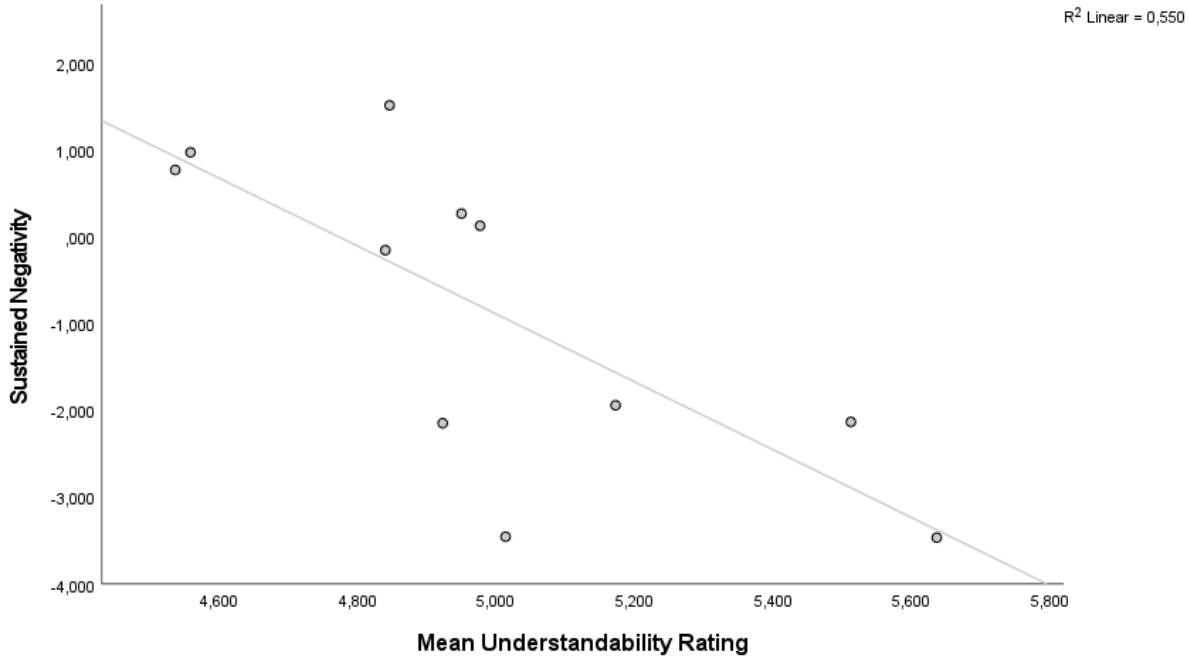


Figure 7

Correlation between mean understandability ratings and sustained negativity measured over frontal sites during evaluation of common uses in the high understandability group



3.2 Creative Idea Evaluation and Mood

3.2.1 Data Preparation

For the analysis of the relationship between mood and evaluation of novelty and understandability, a different dataset was used. For this dataset, the PANAS score, mean novelty rating, and mean understandability rating were obtained. Here it is important to emphasize that these are not the same novelty and understandability ratings as WA participants received for their word associations. These are the mean novelty and understandability ratings that RS participants gave when evaluating the word associations. As RS participants received different associations in different sets, the dataset was split into ten datasets representing the separate sets A through J. On these, correlation tests were performed.

3.2.2 Tests for normal distribution of data

Shapiro Wilk tests indicated that data in all variables were not normally distributed.

3.2.3 Correlations

Spearman's rank-order correlation for evaluation of novelty and understandability with mood of RS participants was computed per set. Between evaluation of understandability and mood of RS participants a significant positive correlation was found in set J (r (6) = .88, $p < .05$) and a significant negative correlation was found in set E (r (5) = -.90, $p < .05$).

4. Discussion

In the study by Rataj et al. (2018a), participants completed an alternate use evaluation task (AUeT) and a word association task (WA). EEG measures of alpha event-related desynchronization and event-related potential obtained during the AUeT, as well as the generated word associations, were made available by the researchers. In this study, the creative ability of this group of WA participants was evaluated. On an understandability scale and a novelty scale, their creative word associations were evaluated by RS participants of the current study. The main aim of this study was to investigate the relationship between novelty and understandability ratings of word associations and alpha event-related desynchronization (ERD) as well as event-related potential (ERP) measured in amplitudes of the N400 and sustained negativity. Furthermore, mood of the RS participants was measured and correlated with evaluation of novelty and understandability, as studies indicate an influence of mood on the evaluation of creativity of others (Mastria et al., 2019). To the best of our knowledge, no research has investigated the relationship between novelty or understandability ratings of word associations and EEG measures recorded during an alternate use evaluation task.

4.1 Creative Ability and EEG Measures

The first important finding, which influences interpretation of all other findings, is the strong negative correlation between novelty and understandability of word associations. As these two concepts measure two sides of the same concept, a correlation was to be expected. The more novel a word association was, the lower its understandability was. This can be explained by the findings that novel word associations usually show large semantic distance, which means that the two associated words are not closely linked in meaning, and hence the association might be harder to understand. Meanwhile, highly understandable associations have small semantic distance but are usually obvious and not novel. Creative associations would be ones that are evaluated as highly novel but still understandable.

In the first hypothesis, novelty ratings of word associations were expected to predict small alpha ERD. This hypothesis had to be rejected as no correlation between novelty and alpha ERD measured over parieto-occipital sites four and eight was found. Also, no correlations between understandability or novelty ratings and alpha ERD measured over parieto-occipital right electrode sites was found. The hypothesis was based on findings of previous studies which indicated an association between alpha power and creative ability (Chrysikou et al., 2021; Kenett et al., 2021; Lustenberger et al. 2018). These studies found

that enhancing alpha power increases the novelty of creative ideas generated by participants, while inhibiting alpha power decreases novelty of creative ideas but increases their appropriateness. As the two-fold model by Kleinmintz et al (2019) suggests that the evaluation phase is involved in the phase of idea generation, a correlation between EEG data measured during the evaluation phase and the result of creative ideation was hypothesized. Hence, it was expected that the ability to create novel ideas would predict smaller alpha ERD. However, this could not be confirmed by the results of this study. However, as the number of participants in the rating-scale study was far below the desired sample size, this conclusion must be interpreted carefully.

The second hypothesis stated that understandability of word associations would predict N400 amplitudes. However, the results of this study found no correlation between N400 amplitudes and understandability. In previous studies, N400 amplitudes were assumed to be associated with conceptual expansion and semantic reintegration (Rataj et al., 2018a; Goldstein, 2012). As the association of two words which are semantically very distant would require semantic reintegration, it was expected that a correlation between N400 amplitudes and the understandability of words would be found. Correlation between N400 amplitudes and novelty would have also been an interesting finding, as novel word associations are most likely ones that have great semantic distance and therefore require semantic integration. However, no correlations between N400 amplitudes and measures of creativity were found. Again, due to the small number of participants in the rating-scale study, this finding does not necessarily have strong implications.

For the third hypothesis, analysis of sustained negativity in the whole sample showed two notable findings. For one, a positive correlation was found between understandability and sustained negativity during evaluation of creative uses of objects and unrelated word pairs. Higher understandability of word associations in the WA task correlated with less sustained negativity measured during evaluation of creative uses of objects and unrelated word pairs. Moreover, simple linear regression results showed that understandability predicts the amplitudes of sustained negativity in these conditions. As second finding, novelty of word associations showed a negative correlation with sustained negativity. Higher novelty of word associations correlated with more negative amplitudes during evaluation of creative uses of objects and unrelated word pairs. Simple linear regression showed that novelty predicts the amplitudes of sustained negativity in these conditions. The study by Rataj et al. (2018a), in which these sustained negativity measures were obtained, found that evaluation of unrelated

word pairs evoked the most sustained negativity. Creative and common use evaluation showed less sustained negativity. In a study by Rutter et al. (2012), where participants evaluated unusualness and appropriateness of literal, novel, and meaningless metaphors, a similar effect was found. In this study, most sustained negativity was measured during the evaluation of meaningless sentences, less with novel sentences, and least during evaluation of literal sentences. The findings of these studies support the recent interpretations of sustained negativity indicating ongoing effort of reinterpretation and increased effort for meaning integration. This idea is further supported by findings from research reported by Jiang et al. (2009). In their research they examined the processing of sentences that contained semantically incongruent information (misused universal quantifiers). Sustained negativity was found in the absence of the N400. This clearly distinguishes the role of sustained negativity from that of the N400 and substantiates the idea that sustained negativity represents a reinterpretation process and the ongoing effort to find meaning after an initial failed attempt. These findings imply that the process of conceptual expansion is continuous as can be interpreted through the N400 and sustained negativity amplitudes.

Regarding the findings of the current study, it can be argued that the ability to generate novel word associations demands conceptual expansion as it requires a meaningful link between two semantically distant words which do not have an obvious connection. This leads to the interpretation that novelty of word associations, as compared to understandability, is the more accurate representation of creative ability in the current study, provided that the generated associations make sense. Considering the importance of evaluation processes during the generation of creative ideas (Kleinmintz et al., 2019), the finding that more novel word associations generated in the WA task predict more sustained negativity during evaluation of creative uses of objects would therefore be in line with the previous findings regarding sustained negativity. The ability to create novel word associations, due to an ability to create a meaningful link between semantically distant words, predicts effort in meaning reintegration, represented by increased sustained negativity, in the evaluation of creative uses of objects of objects. Similarly, the ability to create novel word associations, due to an ability to create a meaningful link between two words, predicts effort in meaning reintegration, represented by sustained negativity, in the identification of unrelated word pairs as such.

Following the interpretation of novelty of word associations indicating high creativity and considering its strong negative correlation with understandability, the most

understandable word associations could be interpreted as the most obvious, so least creative, associations. This would then explain the finding that higher understandability of word associations predicts less sustained negativity. The least creative word associations, for generation of which no meaning reintegration is needed, predict the least effort in meaning reintegration, as represented by less sustained negativity, during evaluation of creative uses of objects and unrelated word pairs. These findings imply that sustained negativity represents resources and effort put into meaning reintegration which is necessary for the evaluation of creative uses of objects and unrelated word pairs and allows for the generation of creative word association.

For the median split, the whole sample was split into group of low and high groups for understandability and novelty. The median split analysis for the novelty scale did not show any significant results. The only significant findings were made in the high and low understandability groups. In the low understandability group, novelty of word associations was found to predict more sustained negativity measured over frontal sites during evaluation of creative uses of objects. This would be in line with the previous findings from the whole sample in which sustained negativity seemed to indicate more effort in meaning reintegration.

Meanwhile, in the high understandability group, it was found that high novelty of word associations predicts less sustained negativity during evaluation of common uses of objects while high understandability predicts more sustained negativity during evaluation of common uses of objects. This relationship goes in the opposite direction of all previous findings in this study. However, the previous findings in the whole sample show a relationship in the opposite direction during the evaluation of creative uses of objects or unrelated word pairs while in this analysis it is the relationship between evaluation of common uses of objects and sustained negativity. As these are findings from the high understandability group, it can be argued that the word associations high in understandability, brought forth by individuals in this group, are the most obvious, so least creative, word associations. This would imply that the least creative word associations predict strong sustained negativity during evaluation of common uses of objects. As these individuals lack the ability to generate creative word associations, it can be assumed that they are not very good at meaning reintegration, meaning that for the evaluation of common uses of objects a lot of effort in meaning reintegration is needed. For individuals in this group that create very novel, so more creative, word associations, less effort is needed during the evaluation of common uses of objects.

Furthermore, in the high-understandability group higher understandability of word associations was found to predict more sustained negativity during evaluation of creative uses of objects. This relationship also goes in the opposite direction of the relationships between understandability ratings and sustained negativity found in the sample without the median split. This would be in line with the findings regarding the relationship between evaluation of common uses and novelty as well as understandability of word associations in the high-understandability group. However, this interpretation would imply that sustained negativity does not represent effort put into meaning reintegration to successfully reintegrate two semantically distant stimuli, as indicated by the previous findings of this study. Instead, sustained negativity would represent the difficulty, or rather the struggle, with meaning reintegration of two semantically distant stimuli. While this could also a be valid interpretation of sustained negativity, these findings must be interpreted very cautiously as they stem from an analysis performed on a much smaller sample than the findings from analysis performed on the whole sample.

Most findings, seem to support the idea that higher creative ability requires more meaning reintegration which is represented by more sustained negativity in the evaluation of semantically distant or incongruent information. Notably, most significant relationships between novelty and understandability rating and sustained negativity were found in the evaluation of creative uses of objects and unrelated word pairs, which leads back to the findings by Rataj et al. (2018a) and Rutter et al. (2012) who found sustained negativity to be strongest in these conditions, compared to evaluation of common, or in their case literal, evaluations.

4.2 Creative Idea Evaluation and Mood

It was expected that a better mood of RS participants would predict higher ratings in novelty and understandability. As previous studies have shown that mood influences the evaluation of creativity of others (Mastria et al., 2019), this relationship was expected in the current study as well. Such a relationship was only found in two of the ten sets. Therefore, it cannot be concluded that mood was found to have any influence on the evaluation of creativity of others. However, as the number of RS participants per set ranged between five and eight the interpretation of these results should not be interpreted as indicative of similar effects in the population.

4.3 Strengths and Limitations

Reflecting on the study and the steps taken to reach these results, a few important remarks are to be made. While the general design of the rating-scale study was very elaborate and took into consideration many factors that could skew the results, there are some limitations in the study design. The design of the scales was well thought out and there were intentionally full descriptions of every point of the seven-point Likert scale shown with every question. Generally, through randomization of the order of word associations, word associations of one WA participant being split across multiple sets, and division into and counterbalancing of two blocks within a set, with a break in between, a lot was done to avoid the design from unintentionally influencing the outcome of the study. However, a point to consider in future research could be the presentation of word associations, their explanations, and the Likert scales. RS Participants based their evaluation of novelty and understandability on the word association and its explanation. However, as the scale to evaluate novelty was presented alongside the explanation for the word association, this might have influenced the perceived novelty of the word association. This could have been prevented by displaying the word association alongside the scale to evaluate novelty first and then the word association and its explanation alongside the scale to evaluate understandability.

Another major limitation is the sample size number of RS participants per set, and at the same time per word association. While many RS participants were recruited for the study (N=123) a majority did not participate in the survey until the end and had to be excluded. As the remaining 57 participants were spread across all sets, no set contained even half the number of intended participants. In this way, every association was evaluated, but the evaluation was performed by a small number of participants. This means that the replies of the individual might have strongly affected the overall results. This effect became even stronger for the analysis of mood as results were compared per set, meaning a sample size of five. This could largely explain the failure to find any correlations between mood and creativity ratings which are found in many other studies. Studies always become more relevant the larger and more representative the sample size is, so this is a common limitation. However, for the aim of the study the relation between RS participants and sets of word associations is notably skewed and can be deemed as influential factor especially for the outcome of the median split analysis and the analysis on mood and creative idea evaluation. As the initial sample size was much larger, but more than half of the RS participants had to be excluded due to not finishing the study, more usable data could have potentially been

obtained by making the study shorter. This, however, would have required even more RS participants to still evaluate the same overall number of word associations. To avoid this issue of sample size, future studies could rely on evaluation of creativity through distributional semantics approaches that utilize semantic vectors.

5. Conclusions and Implications

Findings in the relationship between EEG measures during the evaluation of common object use, creative object use and unrelated word pairs and the ability to create novel and understandable word associations support the recent interpretations of sustained negativity. These findings point to the importance of reinterpretation processes in the phases of creative idea evaluation and creative idea generation. The ability to generate creative word associations was found to predict more effort for meaning construction and semantic reintegration during the evaluation of creative uses of objects and unrelated word pairs. Considering these results, sustained negativity can be interpreted to represent the ongoing effort in meaning reintegration after a failed initial attempt to create a meaningful link between two semantically distant stimuli. This ongoing effort in meaning reintegration seems to be needed in the generation of creative ideas and the evaluation of semantically distant word pairs. Alpha event-related desynchronization, as well as N400 amplitudes measured during the evaluation phase were found not to be predicted by the ability to create novel or understandable word associations. Furthermore, the often-found influence of mood on the evaluation of creativity of others could not be found in this study. These findings need to be interpreted with the strong implications of the small sample size in mind. Especially the analysis of the median split and the relationship between mood and evaluation of creativity in RS participants is heavily impacted by this limitation. This study was the first to investigate the relationship between ability to generate creative ideas, indicated by novelty and understandability ratings of word associations, and EEG measures obtained during an alternate use evaluation task. The findings of this study extend the body of research investigating the relationship between generation and evaluation of ideas and the role of meaning reintegration in this process.

References

Agdam, S. J., & Sadeghi, K. (2014). Two formats of word association tasks: A study of depth of word knowledge. *English Language Teaching*, *7*(10). https://doi.org/10.5539/elt.v7n10p1

Beaty, R. E., Benedek, M., Wilkins, R. W., Jauk, E., Fink, A., Silvia, P. J., Hodges, D. A., Koschutnig, K., & Neubauer, A. C. (2014). Creativity and the default network: A functional connectivity analysis of the creative brain at rest. *Neuropsychologia*, *64*, 92–98. https://doi.org/10.1016/j.neuropsychologia.2014.09.019

Beaty, R. E., Kenett, Y. N., Christensen, A. P., Rosenberg, M. D., Benedek, M., Chen, Q., Fink, A., Qiu, J., Kwapil, T. R., Kane, M. J., & Silvia, P. J. (2018). Robust prediction of individual creative ability from brain functional connectivity. *Proceedings of the National Academy of Sciences*, *115*(5), 1087–1092. https://doi.org/10.1073/pnas.1713532115

Beaty, R. E., Seli, P., & Schacter, D. L. (2019). Network neuroscience of creative cognition: mapping cognitive mechanisms and individual differences in the creative brain. *Current Opinion in Behavioral Sciences*, *27*, 22–30. https://doi.org/10.1016/j.cobeha.2018.08.013

Benedek, M. (2018). Internally directed attention in creative cognition. *The Cambridge Handbook of the Neuroscience of Creativity*, 180–194. https://doi.org/10.1017/9781316556238.011

Benedek, M., Bergner, S., Könen, T., Fink, A., & Neubauer, A. C. (2011). EEG alpha synchronization is related to top-down processing in convergent and divergent thinking. *Neuropsychologia*, *49*(12), 3505–3511. https://doi.org/10.1016/j.neuropsychologia.2011.09.004

Benedek, M., Jauk, E., Sommer, M., Arendasy, M., & Neubauer, A. C. (2014). Intelligence, creativity, and cognitive control: The common and differential involvement of executive functions in intelligence and creativity. *Intelligence*, *46*, 73–83. https://doi.org/10.1016/j.intell.2014.05.007

Benedek, M., Nordtvedt, N., Jauk, E., Koschmieder, C., Pretsch, J., Krammer, G., & Neubauer, A. C. (2016). Assessment of creativity evaluation skills: A psychometric investigation in prospective teachers. *Thinking Skills and Creativity*, *21*, 75–84. https://doi.org/10.1016/j.tsc.2016.05.007

Camarda, A., Salvia, M., Vidal, J., Weil, B., Poirel, N., Houdé, O., Borst, G., & Cassotti, M. (2018). Neural basis of functional fixedness during creative idea generation: An EEG study. *Neuropsychologia*, *118*, 4–12. https://doi.org/10.1016/j.neuropsychologia.2018.03.009

Campbell, D. T. (1960). Blind variation and selective retentions in creative thought as in other knowledge processes. *Psychological Review*, *67*(6), 380–400. https://doi.org/10.1037/h0040373

Ceh, S. M., Edelmann, C., Hofer, G., & Benedek, M. (2021). Assessing raters: What factors predict discernment in novice creativity raters? *The Journal of Creative Behavior*, *56*(1), 41–54. https://doi.org/10.1002/jocb.515

Chrysikou, E. G. (2019). Creativity in and out of (cognitive) control. *Current Opinion in Behavioral Sciences*, *27*, 94–99. https://doi.org/10.1016/j.cobeha.2018.09.014

Chrysikou, E. G., Morrow, H. M., Flohrschutz, A., & Denney, L. (2021). Augmenting ideational fluency in a creativity task across multiple transcranial direct current stimulation montages. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-85804-3

Chrysikou, E. G., Weber, M. J., & Thompson-Schill, S. L. (2014). A matched filter hypothesis for cognitive control. *Neuropsychologia*, *62*, 341–355. https://doi.org/10.1016/j.neuropsychologia.2013.10.021

Cropley, A. (2006). In praise of convergent thinking. *Creativity Research Journal*, *18*(3), 391–404. https://doi.org/10.1207/s15326934crj1803_13

Ellamil, M., Dobson, C., Beeman, M., & Christoff, K. (2012). Evaluative and generative modes of thought during the creative process. *NeuroImage*, *59*(2), 1783–1794. https://doi.org/10.1016/j.neuroimage.2011.08.008

Engelen, U., Peuter, S. D., Victoir, A., Diest, I. V., & van den Bergh, O. (2006). Verdere validering van de positive and negative affect schedule (PANAS) en vergelijking van twee nederlandstalige versies. *Gedrag En Gezondheid*, *34*(2), 61–70. https://doi.org/10.1007/bf03087979

Fink, A., & Benedek, M. (2014). EEG alpha power and creative ideation. *Neuroscience & Biobehavioral Reviews*, *44*, 111–123. https://doi.org/10.1016/j.neubiorev.2012.12.002

Goldstein, A., Arzouan, Y., & Faust, M. (2012). Killing a novel metaphor and reviving a dead one: ERP correlates of metaphor conventionalization. *Brain and Language*, *123*(2), 137–142. https://doi.org/10.1016/j.bandl.2012.09.008

Grohman, M., Wodniecka, Z., & Kłusak, M. (2006). Divergent thinking and evaluation skills: Do they always go together? *The Journal of Creative Behavior*, *40*(2), 125–145. https://doi.org/10.1002/j.2162-6057.2006.tb01269.x

Guilford, J. P. (1956). The structure of intellect. *Psychological Bulletin*, *53*(4), 267–293. https://doi.org/10.1037/h0040755

Guilford, J. P. (1967). Creativity: Yesterday, today and tomorrow. *The Journal of Creative Behavior*, *1*(1), 3–14. https://doi.org/10.1002/j.2162-6057.1967.tb00002.x

Guilford, J. P. (1984). Varieties of divergent production. *The Journal of Creative Behavior*, *18*(1), 1–10. https://doi.org/10.1002/j.2162-6057.1984.tb00984.x

Howard-Jones, P. A. (2002). A dual-state model of creative cognition for supporting strategies that foster creativity in the classroom. *International Journal of Technology and Design Education*, *12*(3), 215–226. https://doi.org/10.1023/a:1020243429353

Jiang, X., Tan, Y., & Zhou, X. (2009). Processing the universal quantifier during sentence comprehension: ERP evidence. *Neuropsychologia*, *47*(8–9), 1799–1815. https://doi.org/10.1016/j.neuropsychologia.2009.02.020

Kalargiros, E. M., & Manning, M. R. (2015). Divergent thinking and brainstorming in perspective: Implications for organization change and innovation. *Research in Organizational Change and Development*, 293–327. https://doi.org/10.1108/s0897-301620150000023007

Kenett, Y. N., Anaki, D., & Faust, M. (2014). Investigating the structure of semantic networks in low and high creative persons. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00407

Kenett, Y. N., Gold, R., & Faust, M. (2018). Metaphor comprehension in low and high creative individuals. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00482

Kenett, Y. N., Rosen, D. S., Tamez, E. R., & Thompson-Schill, S. L. (2021). Noninvasive brain stimulation to lateral prefrontal cortex alters the novelty of creative idea generation. *Cognitive, Affective, & Behavioral Neuroscience*, *21*(2), 311–326. https://doi.org/10.3758/s13415-021-00869-x

Kleinmintz, O. M., Ivancovsky, T., & Shamay-Tsoory, S. G. (2019). The two-fold model of creativity: The neural underpinnings of the generation and evaluation of creative ideas. *Current Opinion in Behavioral Sciences*, *27*, 131–138. https://doi.org/10.1016/j.cobeha.2018.11.004

Luft, C. D. B., Zioga, I., Thompson, N. M., Banissy, M. J., & Bhattacharya, J. (2018). Right temporal alpha oscillations as a neural mechanism for inhibiting obvious associations. *Proceedings of the National Academy of Sciences*, *115*(52). https://doi.org/10.1073/pnas.1811465115

Lustenberger, C., Boyle, M. R., Foulser, A. A., Mellin, J. M., & Fröhlich, F. (2015). Functional role of frontal alpha oscillations in creativity. *Cortex*, *67*, 74–82. https://doi.org/10.1016/j.cortex.2015.03.012

Mastria, S., Agnoli, S., & Corazza, G. E. (2019). How does emotion influence the creativity evaluation of exogenous alternative ideas? *PLoS ONE*, *14*(7). https://doi.org/10.1371/journal.pone.0219298

Miller, E. K., & Cohen, J. D. (2001). An Integrative Theory of Prefrontal Cortex Function. *Annual Review of Neuroscience*, *24*(1), 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

Ni, M., Yang, L., Chen, J., Chen, H., & Li, X. (2014). How to improve divergent thinking capability by information technology and extenics. *Procedia Computer Science*, *31*, 158–164. https://doi.org/10.1016/j.procs.2014.05.256

Nijstad, B. A., de Dreu, C. K. W., Rietzschel, E. F., & Baas, M. (2010). The dual pathway to creativity model: Creative ideation as a function of flexibility and persistence. *European Review of Social Psychology*, *21*(1), 34–77. https://doi.org/10.1080/10463281003765323

Peeters, F.P.M.L., Ponds, R.W.H.M., & Vermeeren, M.T.G. (1996). Affectiviteit en zelfbeoordeling van depressie en angst. Tijdschrift voor Psychiatrie, 38, 240-250.

Rataj, K., Nazareth, D. S., & van der Velde, F. (2018a). Use a spoon as a spade?: Changes in the upper and lower alpha bands in evaluating alternate object use. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01941

Rataj, K., Przekoracka-Krawczyk, A., & van der Lubbe, R. H. (2018b). On understanding creative language: The late positive complex and novel metaphor comprehension. *Brain Research*, *1678*, 231–244. https://doi.org/10.1016/j.brainres.2017.10.030

Runco, M. A., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal*, *24*(1), 92–96. https://doi.org/10.1080/10400419.2012.650092

Rutter, B., Kröger, S., Hill, H., Windmann, S., Hermann, C., & Abraham, A. (2012). Can clouds dance? Part 2: An ERP investigation of passive conceptual expansion. *Brain and Cognition*, *80*(3), 301–310. https://doi.org/10.1016/j.bandc.2012.08.003

Silvia, P. J. (2008). Discernment and creativity: How well can people identify their most creative ideas? *Psychology of Aesthetics, Creativity, and the Arts*, *2*(3), 139–146. https://doi.org/10.1037/1931-3896.2.3.139

Smallwood, J., Bernhardt, B. C., Leech, R., Bzdok, D., Jefferies, E., & Margulies, D. S. (2021). The default mode network in cognition: A topographical perspective. *Nature Reviews Neuroscience*, *22*(8), 503–513. https://doi.org/10.1038/s41583-021-00474-4

Sowden, P. T., Pringle, A., & Gabora, L. (2014). The shifting sands of creative thinking: Connections to dual-process theory. *Thinking & Reasoning*, *21*(1), 40–60. https://doi.org/10.1080/13546783.2014.885464

Vartanian, O., Beatty, E. L., Smith, I., Forbes, S., Rice, E., & Crocker, J. (2019). Measurement matters: The relationship between methods of scoring the alternate uses task and brain activation. *Current Opinion in Behavioral Sciences*, *27*, 109–115. https://doi.org/10.1016/j.cobeha.2018.10.012

Ward, T. (1994). Structured imagination: The role of category structure in exemplar generation. *Cognitive Psychology*, *27*(1), 1–40. https://doi.org/10.1006/cogp.1994.1010

Appendix A

Filler Items included in each set to include the high understandability and high novelty end of the scales

Word	Association	Explanation	Filler type
Benzine	Auto	Een auto verbruikt benzine	understandable
Verjaardag	Cadeautjes	Er zijn cadeautjes op de verjaardag	understandable
Gezicht	Ogen	Ogen maken deel uit van je gezicht	understandable
Kantoor	Bureau	In een kantoor staan bureau's	understandable
Kat	Dier	Een kat is een dier	understandable
Kussen	Bed	Kussens liggen op een bed	understandable
Schelp	Stoel	Je kan er mee klimmen	original
Lantaarn	Afzuigkap	Beide zijn erg snel	original
Stopcontact	Sloopkogel	Omdat het stroomt	original
Tafel	Purpur	Tafelpoten zijn vaak rond	original
Palmboom	Astronaut	Kokosnoten zijn bruin	original
Steen	Grap	Ik hou van stenen.	original

Appendix B

Items of the Dutch adaption by Peeters et al. (1996) of the PANAS mood scale

Panas item	Versie van Peeters et al. (1996)
1	Geïnteresseerd
2	Uitgelaten
3	Sterk
4	Enthousiast
5	Trots
6	Alert
7	Geïnspireerd
8	Vastberaden
9	Aandachtig
10	Actief

Appendix C

Instructions shown to RS participants before beginning the evaluation

Instructies

U krijgt straks woordparen te zien die een associatie hebben met elkaar.

Voor elk woordpaar is het 1e woord, bijvoorbeeld '**voeten**', een woord waarvoor een associatie is gemaakt, bijvoorbeeld '**schoenen**'. Deze woorden zullen in het onderzoek als volgt worden weergegeven: **Voeten → Schoenen.** In het onderzoek zal er een uitleg staan van de associatie onder elk woordpaar: **Voeten → Schoenen.**

Uitleg: Want schoenen draag je aan je voeten.

In dit voorbeeld wordt voor het woord 'Voeten' (1e woord) de associatie 'Schoenen' (2e woord) gegeven, met als uitleg dat u schoenen aan uw voeten draagt.

Een ander voorbeeld is: **Sleutel → Brieven**

Uitleg: Je kan je sleutel gebruiken om brieven te openen, zoals een briefopener.

In dit voorbeeld wordt voor het woord 'Sleutel' (1e woord) de associatie 'Brieven' (2e woord) gegeven, met als uitleg dat u een sleutel kunt gebruiken als een soort briefopener om zo brieven te openen.

Nog een voorbeeld is: **Fiets → Komma**

Uitleg: Beiden zijn zwart.

In dit voorbeeld wordt voor het woord 'Fiets' (1e woord) de associatie 'Komma' (2e woord) gegeven, met als uitleg dat beiden zwart zijn.

Appendix D

Dataset with alpha desynchronization and mean creativity scores of WA participants

Participant Number	Novelty Mean	Understandability Mean	PO4 Common	PO4 Creative	PO4 Anomalous	PO8 Common	PO8 Creative	PO8 Anomalous
1	4.72	3.60	-63.61	-69.16	-63.49	-70.40	-71.59	-69.52
2	5.14	2.95	-55.38	-14.34	-32.78	-41.70	3.64	-14.40
3	3.49	5.64	-23.02	-26.07	-27.88	-24.15	-32.14	-17.18
4	3.86	4.51	-15.75	-16.57	-25.10	-51.31	-49.24	-62.86
5	3.75	5.01	-61.30	-46.46	-50.82	-78.57	-74.71	-73.45
6	4.05	3.69	-14.77	-28.86	5.97	-53.31	-52.52	-21.12
7	4.62	4.49	-35.93	-49.06	-46.73	-83.45	-81.06	-82.34
8	4.40	4.85	-60.91	-55.60	-62.46	-69.06	-58.08	-68.53
9	4.64	3.86	-72.59	-43.11	-51.56	-70.63	-55.67	-47.20
10	3.58	5.17	-18.09	-5.17	-4.46	-6.56	-16.51	-12.02
11	3.99	4.84	-67.47	-40.05	-52.49	-86.97	-73.08	-80.59
12	4.34	4.56	-9.04	14.82	9.00	-42.69	-16.87	-18.89
13	2.99	5.51	-34.26	-38.38	-4.88	-30.25	-26.44	-12.14
14	4.60	4.54	-16.93	-2.97	-10.75	-47.83	-29.56	-40.40
15	3.97	4.95	-33.83	14.28	0.00	-45.60	-21.33	-32.30
16	5.86	3.05	-52.54	-62.81	-23.71	-65.08	-75.01	-27.31
17	3.94	4.92	-44.55	-3.42	-36.15	-25.60	3.43	-4.41
18	4.45	4.36	41.01	155.60	56.33	-8.54	48.51	-4.24
19	3.63	4.98	-65.20	-41.22	-58.55	-66.79	-52.04	-67.15
20	4.15	4.46	-76.23	-62.45	-55.91	-80.07	-76.13	-71.78
21	5.63	3.18	-20.37	52.79	53.87	-41.33	-15.52	-24.89
22	5.73	3.30	-44.05	-58.25	-42.71	-67.34	-73.49	-67.05

Appendix E

Dataset with N400 amplitudes and mean creativity scores of WA participants

Participant Number	Novelty Mean	Understandability Mean	Frontal Common	Frontal Creative	Frontal Anomalous	Central Common	Central Creative	Central Anomalous
1	4.72	3.60	-0.40	-1.81	-2.59	-5.50	-5.85	-6.10
2	5.14	2.95	-8.28	-10.33	-8.73	-6.45	-6.70	-7.03
3	3.49	5.64	-3.59	-2.78	-4.64	-1.88	-0.99	-2.36
4	3.86	4.51	0.71	0.24	-0.22	-2.55	-2.55	-3.62
5	3.75	5.01	-5.02	-3.74	-4.29	-5.76	-5.20	-4.30
6	4.05	3.69	0.59	-0.61	-0.66	-0.91	-0.59	-1.59
7	4.62	4.49	-3.75	-4.44	-4.30	-3.27	-3.42	-2.76
8	4.40	4.85	0.00	-3.09	-4.24	-2.67	-4.76	-3.39
9	4.64	3.86	-1.38	0.18	-3.15	-5.13	-4.19	-4.89
10	3.58	5.17	-2.42	-2.03	-2.37	-1.89	-1.91	-1.93
11	3.99	4.84	-3.19	-1.92	-2.75	-2.47	-2.93	-3.42
12	4.34	4.56	-2.67	-0.54	-1.16	-1.24	-3.15	-4.01
13	2.99	5.51	-0.11	-0.51	0.74	-2.81	-2.68	-2.86
14	4.60	4.54	-1.35	-0.79	-0.48	-1.27	-1.05	-0.39
15	3.97	4.95	-1.31	-2.36	-1.48	-1.42	-3.30	-2.32
16	5.86	3.05	0.46	-1.24	-0.22	-0.43	-2.73	-1.09
17	3.94	4.92	-3.25	-3.03	-3.16	-0.54	-0.62	-1.03
18	4.45	4.36	-0.22	0.78	0.27	0.01	-1.45	-1.16
19	3.63	4.98	-0.58	-1.59	-2.21	-0.81	-1.63	-2.38
20	4.15	4.46	-1.15	-1.70	-1.51	-3.51	-2.44	-2.34
21	5.63	3.18	-3.01	-2.74	-4.18	-3.22	-4.28	-5.33
22	5.73	3.30	0.14	-2.28	-1.85	-2.69	-2.82	-2.70

Appendix F

Dataset with sustained negativity and mean creativity scores of WA participants

Participant Number	Novelty Mean	Understandability Mean	Frontal Common	Frontal Creative	Frontal Anomalous	Central Common	Central Creative	Central Anomalous
1	4.72	3.60	-1.98	-2.69	-4.09	-4.09	-4.73	-4.20
2	5.14	2.95	-2.29	-4.03	-5.21	-1.71	-2.09	-2.85
3	3.49	5.64	-3.47	-3.47	-4.20	-0.60	0.31	-0.51
4	3.86	4.51	1.31	-0.72	-1.26	-1.88	-1.43	-2.08
5	3.75	5.01	-3.46	-3.75	-5.12	-4.03	-3.75	-2.39
6	4.05	3.69	-1.76	-1.40	-2.46	-0.85	-0.45	-1.86
7	4.62	4.49	-3.07	-3.13	-3.40	-2.05	-2.26	-0.70
8	4.40	4.85	1.50	-0.62	-2.11	-1.70	-3.04	-2.54
9	4.64	3.86	-1.01	-0.33	-2.77	-3.48	-3.37	-3.13
10	3.58	5.17	-1.95	-2.36	-3.74	-0.43	-1.26	-1.71
11	3.99	4.84	-0.16	1.26	-0.10	-0.76	-0.03	-0.74
12	4.34	4.56	0.96	1.04	-0.54	1.05	-0.72	-2.18
13	2.99	5.51	-2.14	-1.22	0.16	0.63	0.32	-0.33
14	4.60	4.54	0.76	0.77	0.60	-1.17	-1.15	0.42
15	3.97	4.95	0.26	0.62	-1.59	0.53	-0.98	-0.58
16	5.86	3.05	1.84	-0.62	-0.47	-0.83	-2.80	-1.75
17	3.94	4.92	-2.15	-1.79	-1.99	0.41	0.58	0.38
18	4.45	4.36	0.10	1.14	1.09	-0.73	-1.15	-1.84
19	3.63	4.98	0.12	-0.16	-1.70	-0.51	-0.63	-1.47
20	4.15	4.46	-1.05	-1.08	0.51	-3.40	-1.35	-2.28
21	5.63	3.18	-0.98	-1.39	-2.79	-2.03	-3.67	-3.62
22	5.73	3.30	-1.23	-2.56	-2.68	-3.75	-3.37	-3.19

Appendix G

Dataset of mood scores and mean novelty and understandability rating in Set A.

Participant No.	Mood Score	Novelty Mean	Understandability Mean
1	16	3.92	5.28
2	20	3.79	3.18
3	23	4.26	3.64
4	10	4.00	4.51
5	-7	5.54	5.00
6	3	4.03	4.87