UNIVERSITETET
I OSLO

Masteroppgave

I’ll give you a hint

Examining curiosity-driven exploratory behaviour in a novel non-instrumental information sampling paradigm

Per Flatebø

Professional degree in Clinical Psychology

Number of credits: 30

Departement of Psychology

Faculty of Social Sciences, University of Oslo

Summary:

Author: Per Flatebø

Title: I'll give you a hint: Examining curiosity-driven explorative behaviour in a novel non-instrumental information sampling paradigm

Supervisor: Thomas Espeseth

Research question: Do we adhere to the principle of allocating the least effort when engaging in non-instrumental information sampling, or are we inclined to surpass that threshold by investing extra resources to experience a sense of agency? Can the initial feelings of certainty, curiosity, or the subsequent aha-experience predict exploratory behaviour?

Background: In this study, I examine curiosity-driven exploratory behaviour. First, I investigate the conditions under which curiosity is evoked, from the *information gap*-perspective. Subsequently, I examine two competing models regarding curiosity-driven exploration. In particular, the *expected information gain* (EIG)-model propose that the potential informational gain associated with novel information is the motivational factor driving non-instrumental exploration. A cost-benefit analysis of the predicted costs and reward will ensure no wasted resources. Only when the net value of the evaluation is positive, exploratory behaviour will occur. The competing model, the *agency*-model, states that the act of investing resources is rewarding, as a feeling of self-efficacy follows from exploration. Thus, exploratory behaviour may precede the threshold of the least required effort. Third, I investigate the relation between exploratory behaviour and the subsequent aha-experience in response to curiosity-relief. I propose the aha-experience to be a function of curiosity, but to also be influenced by exploratory behaviour.

Method: To investigate curiosity-driven exploration, a new experimental paradigm was developed: Participants were placed in a situation in which various levels of uncertainty was evoked with Mooney Images (MI). MIs are grey-scale images (GSI) which have been subjected to a modification routine in which blurring and then thresholding results in a seemingly randomized set of white and black spots. This distorts the content of the image causing uncertainty, which eventually can cause curiosity about the pre-modification content. Uncertainty evoked by MIs has earlier been shown to induce curiosity, and releasing the uncertainty by revealing the corresponding GSI has been shown to cause an aha-experience. To contrast the agency and EIG-model participants were asked to choose one out of two ways to obtain the content of the GSI: they could request to see the GSI without any additional

work. The other way was to indulge in exploratory behaviour by requesting visual hints to uncover the GSI content. Curiosity, certainty, and aha-experience was measured with self-report on visual analogue rating scales. A newly developed ‘objective’ calculation of certainty was also included.
Results: Certainty did show a monotonic, negative correlation with curiosity. Exploratory behaviour was positively predicted by curiosity and negatively predicted by certainty, lending support to the agency-model. Aha-experiences were shown to be higher for participants who chose to explore. It was also positively correlated with curiosity, and this relation was moderated by exploration.
Conclusion: The results suggest that exploration is not just an essential means for individuals to access valuable information, but to be inherently rewarding. Additionally, the process of acquiring new information is revealed to play a role in the subsequent aha-experience.

“Because it’s there”

- George Mallory.

Acknowledgement

Thanks to friend and family who has supported me throughout this project. I would specifically thank Kristine. I would also like to thank my supervisor, Thomas, for support and guidance throughout the process.

Oslo, October 2023

TABLE OF CONTENTS

TABLE OF FIGURES

1 INTRODUCTION

Acquisition and processing of knowledge is an inherent ability of all animals, and an indispensable skill in almost every thinkable aspect of life. Without the ability to acquire new knowledge it is hard to imagine any life at all. However, while our environment is abundant with information which is available for sampling, information-processing abilities of human and non-human animals are severely limited. Thus, an organism needs the ability to choose between all information it has the capacity of obtaining through its senses or from internal processes, when engaging in information sampling. Therefore, we need adaptive and powerful mechanisms for selecting and sampling information wisely.

In natural environments, curiosity-driven exploration is a central mode of non-instrumental information sampling. Curiosity, defined as the desire to know, is non-strategic in the sense that the motivation driving the information sampling goes beyond obtaining instrumental rewards, or preserving purely hedonic motivations (Gottlieb & Oudeyer, 2018; Sharot & Sunstein, 2020; Wang & Hayden, 2021). Curiosity a basal component of our cognition and is described is manifested in art and science throughout human history (Loewenstein, 1994). Curiosity has been shown to vary across individuals (Silvia & Christensen, 2020), and to be associated with well-being (Gallagher & Lopez, 2007).

From a temporal perspective, information sampling evoked by curiosity could be framed as a three-stage process. See Figure 1.1. Firstly, a feeling of uncertainty about a particular piece of information must occur. The experience of uncertainty elicits a metacognitive evaluation which may or may not result in a state of curiosity. Secondly, based on the level of experienced curiosity, a decision of whether to investing cognitive or physical resources to sample new information will be taken. Thirdly, an experiential component (an aha-experience) based on the actual information gain, and potentially the resource allocation involved in sampling information will occur.

Here, I will investigate the three stages of curiosity-driven behaviour, from the initial uncertainty to the subsequent aha-experience in a novel experimental paradigm. The experimental procedure takes advantage of stimulus material called Mooney Images (MI). MIs are created by using grey scale images (GSI) which is blurred and then thresholded. The effect of this procedure is that the original content is obstructed and difficult to recognize, sometimes to the extent that is unrecognizable without hints. The result looks like a random distribution of black and white patches which has little resemblance to the GSI from which it was forged. Looking at a MI followed by its corresponding GSI has been found to evoke curiosity (about the content of the MI), followed by an aha-experience (when the content is revealed) (Van de Cruys et al., 2021). Expanding on this research, an exploration condition which allowed participants to request hints (partial revealing the GSI content) was included, enabling investigation of the exploratory nature of curiosity. During the experimental procedure the participants were asked to rate their curiosity and certainty, after viewing the MIs. After having seen the GSI content, the participants were asked to rate their aha-experience (see Figure 2.3. below for an illustration of the experimental procedure). An objective calculation of certainty, called *semantic entropy*, based on the distribution of guesses across all participants were also included.

Most theories concerned with the relation between uncertainty and curiosity share some common concepts and agree that curiosity is a metacognitive feeling. For curiosity to occur, a meta-cognitive function identifying an *informational gap* between the knowledge base of an observer and a new information is necessary (Loewenstein, 1994). An informational gap will be proportional with experienced uncertainty. This metacognitive evaluation compares new informational input to what is already known to the observer and estimates the "size" of the information gap. Most theories conceptualize curiosity as a prediction of the reward associated with filling such an informational gap. Put shortly, a metacognitive evaluation, not only of the information gap, but also physical and cognitive resources available to 'fill' the perceived information gap, must occur for a feeling of curiosity to appear. Thus, if learnable information is to be prioritized, only information gaps that are believed to be bridgeable should lead to the feeling of curiosity and potentially active exploration (Wade & Kidd, 2019). The exact characteristics of an informational gap for curiosity to be experienced is still under debate. However, the theories often align with one out of two main branches of theories called the *novelty*-model and the *prediction error*-model. The novelty perspective states that a linear negative correlation is to be found between the magnitude of an information gap and curiosity. The prediction error perspective predicts intermediate information gaps to evoke the most curiosity, thus, a quadratic model to better explain the relation.

The metacognitive evaluation of information gaps enables the two subsequent steps towards eventual exploratory behaviour: 1) by identifying an information gap the observer can estimate the probability that a new informational input carries the correct properties to fill this gap, and 2) whether the observer has the cognitive and physical abilities to bridge the information gap.

The second stage in the model addresses the relation between curiosity and exploratory behaviour. In what way does curiosity drive exploration? I will contrast the *expected information gain* (EIG)-model, and the *agency*-model when investigating these questions. So far, only one study has investigated the contrasting models using trivia questions. Here, an alternative approach to the trivia questions is provided. One of the most common conceptualizations of curiosity is through the EIG-model. When confronted with an informational input, according to this conceptualization, an individual will evaluate the EIG associated with pursuing that certain informational input. If the EIG is evaluated to exceed the effort allocation which is predicted to be needed to obtain EIG, the individual will allocate resources to the relevant task. This conceptualization of curiosity is supported by evidence showing that infants are more attentive to artificial grammar with learnable patterns, as opposed to grammar where no underlying structures is to be found, showing that children spent their mental resources on informational inputs that conveys a learnable pattern (Gerken et al., 2011). Infants also have been found to prefer patterns of intermediate predictability, when selectively paying attention to auditory or visual information, again, preventing resources from being wasted on information which is too easy, or information which is too complex (Kidd et al., 2014). Studies on adults shows preference for intermediate complexity both when it comes to artworks or musical chords (Day, 1981; Witek et al., 2023), suggesting a stable innate preference for informational input of intermediate complexity or predictability throughout the life span. This perspective follows the trajectories of economic models where human resource allocation in non-instrumental decision-making never exceeds the absolute lowest required resources, to preserve physical and cognitive resources. One such model is the *motivational intensity theory* (MIT) (Brehm & Self, 1989, see Silvestrini et al. (2022) for a in depth review of MIT and more recent motivational theories).

Recently, however, a novel perspective on the motivational properties of curiosity has been communicated through the *agency*-model. According to the agency-model, information sampling may be valuable to observers beyond information gain per se (Metcalfe et al., 2021)

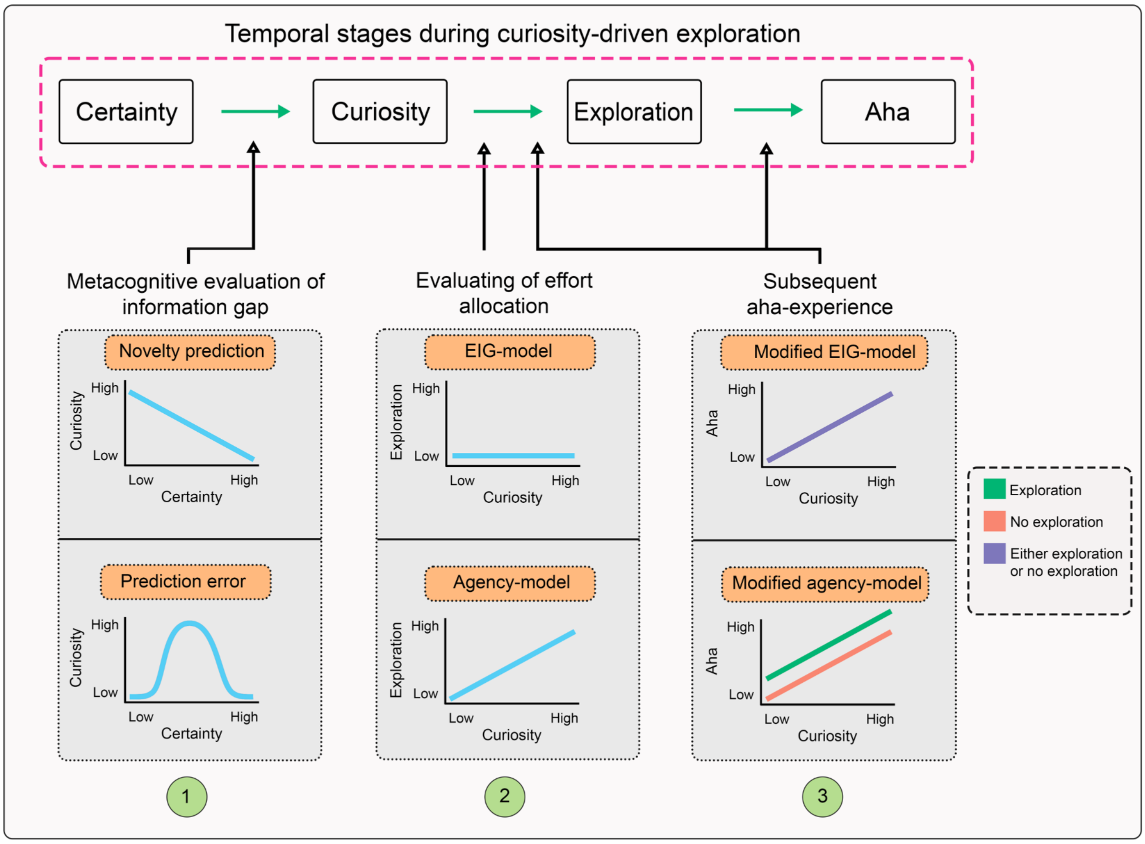
For example, the activity of information sampling may itself be valued by the individual because it may promote feelings of self-efficacy (Clark, 2018; Metcalfe et al., 2021). Some recent evidence has proven more compatible with the agency-model for curiosity-driven behaviour, than with the EIG-perspective (Marvin & Shohamy, 2016; Metcalfe et al., 2021).

Researchers have reasoned that if our cognitive evaluation system utilize the degree of phenomenal curiosity to evaluate the EIG of informational inputs from our environments, there should also be a cognitive function measuring the *actual* information gain. One explanation that has been suggested for the experiential component of the actual information gain is in the subsequent *Aha-Erlebnis* (Kounios & Beeman, 2014; Van de Cruys et al., 2021). The *Aha-Erlebnis* (referred to as the *aha-experience* hereafter) is a sudden realization of increased insight. It has been linked to aesthetic appreciation (Muth & Carbon, 2013) and the process of going from a state of high uncertainty to lower uncertainty (Van de Cruys, 2017). Most research on aha-experience is concerned with passively evaluating stimuli material, and then having participants rate their aha-experience. Some relevant experiments trying to measure the aha-experience by varying complexity (Terwilliger, 1963) or the entropy of a stimulus material (Van de Cruys et al., 2021), or measuring self-reported curiosity or certainty and its effects on aha-experiences (Terwilliger, 1963; Van de Cruys et al., 2021) has been interpreted as evidence that aha-experiences track insight moments.

Although evidence from aha-research has been interpreted as support for the assumption that aha-experiences tracks the actual information gain, another potentially significant determinant for aha-experiences could still have been left out of the equation. One could speculate the reported aha-experience to not merely reflect the actual insight per se, but to also somehow reflect the degree of agency in the sampling process. No studies have yet looked explicitly into how aha-experiences are formed by prior effort allocation. However, one could speculate that the observed willingness to work for information for which curiosity has been reported (Metcalfe et al., 2021), is motivated by a striving for an increase in the subsequent aha-experience. Here, the aha-experience has been incorporated into modified versions of the agency and EIG-model, as there has been no model explicitly addressing the relationship between curiosity, cognitive effort, and aha-experiences (see Figure 1.1).

In the following I aim to investigate the three stages of curiosity-driven exploration proposed initially through a new non-instrumental experimental paradigm.

Figure 1. 1 Illustration of three stages of curiosity-driven exploration



Note: Model for the different stages involved in curiosity-driven exploratory behaviours. Within the dashed, pink line the four stages are shown from a temporal perspective, starting from the left, following the arrows to the right. The green arrows indicate a relationship where the concept on the left side of the arrow is proposed to predict the concept on the right-side of the arrow. The black arrows pointing towards the green arrows shows which models are proposed to explain the relations. Plots for each model shows the relationship as predicted by each model. 1) The relationship between certainty and curiosity has been explained both by the novelty theory and the prediction error theory. 2) Predictions of the relationship from the EIG- and agency-model. Note that the predictions shown here points to situations in which a free choice between exploration or receiving information 'free of charge' is to be made. That is, the exploration is independent of the information gain, which is a certainty. 3) Two models which, extending from previously established models, predicts different outcomes on the subsequent aha-experience. The modified EIG-model predicts a linear relationship between curiosity and aha-experiences independent of exploration, whereas the modified agency-model shows exploration to elevate the aha-experience.

1.1 Navigating our environment in the absence of external rewards

Humans are under the constant pressure of prioritizing certain informational inputs in our surroundings as more valuable than others, due to limitations in our cognitive processing capacities. Therefore, information-sampling policies optimizing selection of relevant information is crucial. In neuropsychological literature, a distinction is usually drawn between sampling policies based on the instrumental utility of the information being sampled

(Gottlieb et al., 2020). This distinction discriminates based on whether the information can be exploited for external rewards or not (Dubey et al., 2022). Material goods, increased social reputation, or safety in the foreseeable future, resulting from obtaining a specific piece of information, are examples of reward associated with *instrumental information sampling* according to this theoretical framework. When an organism engages in what is referred to as *non-instrumental information sampling*, the behaviour does not have any apparent instrumental function, and is driven by internal reward systems (Ten et al., 2021). Curiosity is often seen as a type of non-instrumental information sampling (Gottlieb & Oudeyer, 2018; Loewenstein, 1994; Metcalfe et al., 2021), with behaviour facilitating the development of mental maps being another example of non-instrumental information sampling (Tolman, 1948; Wang & Hayden, 2021).

1.2 Information gaps

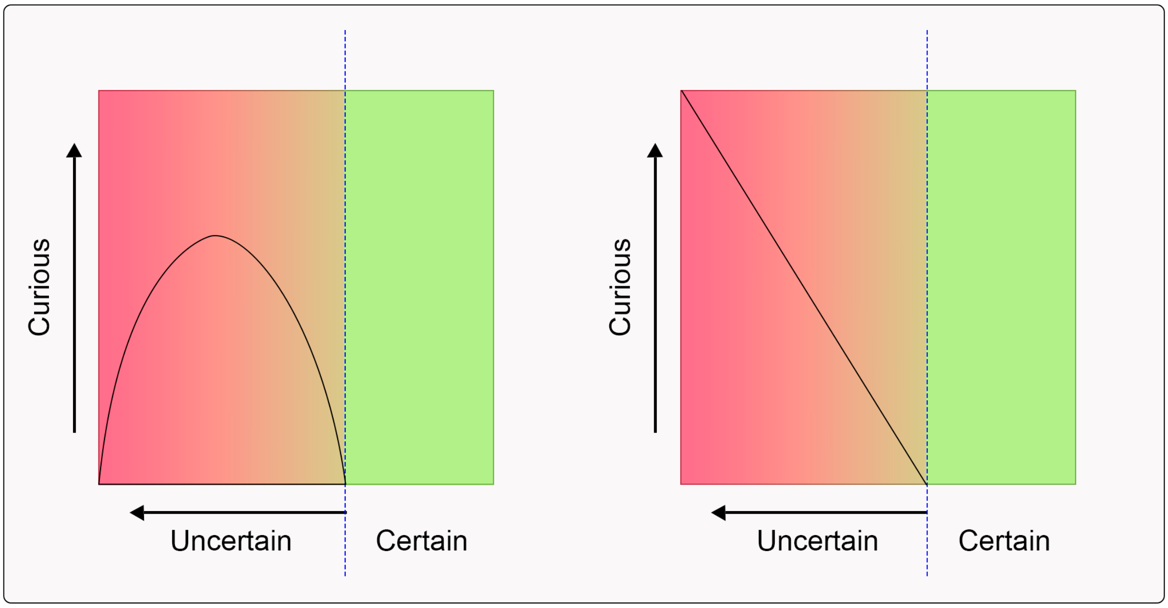
In the movie Pulp Fiction, when introducing the character Marsellus Wallace, careful considerations are taken to prevent the viewer from seeing his face, by only filming the back of his head, his footsteps, or blurring parts of the movie frame where the character appears. This causes an interesting situation in which the viewer has access to almost all the information required for developing an understanding of the character, including visual appearance, his way of expressing himself verbally, and his interaction with other characters, while still highlighting the absence of the most identifying feature of a human being: its face. This storytelling technique forces a metacognitive realization onto the viewer: the knowledge about the absence of knowledge. The metacognitive realization is caused by elevating the interpretative freedom of the viewer by introducing an unknown variable, the hidden face, elevating the level of entropy for the information. In the field of curiosity research, this is referred to as an information gap (Loewenstein, 1994).

Most current researchers converge on the fact that curiosity is a meta-cognitive feeling which can be understood from the information gap perspective (Gottlieb & Oudeyer, 2018; Loewenstein, 1994; Son & Metcalfe, 2000; Van de Cruys et al., 2021). This perspective on curiosity states that metacognitive abilities which enables evaluation of what one already knows, and what one does not know, is obligatory prerequisites for experiential curiosity. Importantly, this does not mean that all information gaps lead to curiosity. It is obvious that the amount of knowledge which is obtainable for even the most competent humans is miniscule, considering that the pool of all knowledge in theory is infinite. This means that the information which is unknown (information gaps) represent more information than what is known. However, most information gaps do not cause curiosity, there is no guarantee that the detection of an information gap causes curiosity, as it depends on the gap being judged as ‘bridgeable’ with current cognitive and physical resources. In other words, the gap has to be framed as a problem which can be solved, given that effort is invested (Wade & Kidd, 2019).

1.2.1 Uncertainty

Experiencing a gap in knowledge is pivotal for the curiosity sensation and potential exploratory behaviour, but there is no consensus when it comes to how the magnitude of uncertainty evoked by an information gap and the curiosity experience is evaluated. Among numerous theories addressing the relation between curiosity and certainty, two “main branches” of theories stand out, as most theories do seem to align with on or the other: the novelty perspective states that a monotonic increasing relation between certainty and curiosity better explain the relationship, while the prediction error perspective predicts a quadratic relationship (Berlyne, 1954; Dubey & Griffiths, 2020; Loewenstein, 1994). The theories are shown in Figure 1.2. Evidence is, however, not unequivocal in support of either of the theories. Some evidence is showing support for the prediction error-explanation (Baranes et al., 2015; Kang et al., 2009; Marvin & Shohamy, 2016; Son & Metcalfe, 2000), while other studies is leaning more towards the novelty theory (Van de Cruys et al., 2021; Van Lieshout et al., 2018).

Figure 1. 2 Relationship between uncertainty and curiosity



Note: The two broad categories in which most theories of the relationship between uncertainty and curiosity can be sorted are the prediction error (illustration to the left) and the novelty-based categories (illustration to the right). The prediction error-model of theories says that curiosity is highest when predictions keep improving. The novelty-model propose a positive monotonic relationship between uncertainty and curiosity.

1.3 Exploratory behaviour

Curiosity creates a motivational incentive to reduce the information gap which oftentimes result in exploratory behaviour (Marvin & Shohamy, 2016; Metcalfe et al., 2021). Since engaging in cognitive or physical labour often is a necessity when obtaining novel information, it is often viewed as a cost which can only be justified by the reward associated with the certainty reduction. However, recent studies have come up with research designs enabling new perspective on the link between curiosity and resource allocation.

1.3.1 The expected information gain-model

Curiosity has been described, by philosophers and psychologists alike, as an “appetite” or “thirst” for knowledge (Loewenstein, 1994). Borrowing terminology which describes unpleasant states motivating fulfilment of physiological needs to describe curiosity seems reasonable, given the observable behaviour in both instances aim at reducing the magnitude of an internal state, hence the similarity between hunger-driven behaviour, and curiosity-driven behaviour (Marvin & Shohamy, 2016). This observation has led researchers to speculate that the same underlying neural networks are motivating goal directed behaviour, independent of the reward being internal or external. Studies have shown an association between curiosity driven behaviour and dopaminergic frontal activation (Bromberg-Martin & Hikosaka, 2009; Kang et al., 2009; Lau et al., 2020). Similar activation patterns are associated with obtaining primary reward in reinforcement learning paradigms, as well as experiencing prediction errors in the case of external motivation (Bayer & Glimcher, 2005; Daw & Doya, 2006; O'Doherty et al., 2002). These findings are often interpreted as supporting the EIG-model, i.e. the primary motivational goal when engaging in curiosity-driven behaviour is to obtain the reward. Earlier studies have shown participants to be willing to engage effortfully with a demanding task in numerous ways in order to obtain information, including opportunity cost (waiting time) (Marvin & Shohamy, 2016), hint-seeking for trivia questions (Metcalfe et al., 2021) and physical effort allocation (Goh et al., 2021).

1.3.2 The agency-model

However, the same neural networks have been shown to be involved in agentic decision-making. Activation in dopamine/striatal areas of the brain when receiving a reward as a consequence of autonomous decision-making, as opposed to passively receiving a reward was associated with enhanced activation (Leotti & Delgado, 2011, 2014; Murayama et al., 2015; Tricomi et al., 2004). Also, it has been argued that dopamine is related to 'wanting', rather than 'liking' or learning (Berridge & Kringelbach, 2015). Therefore, merely wanting a particular piece of information, i.e., just being curious, could explain the dopaminergic activation rather than receiving the rewarding information.

Furthermore, research have shown a preference for being in control, as opposed to taking a more passive role when given the opportunity (Bjork & Hommer, 2007; Markant & Gureckis, 2014; Sharot et al., 2010). Participants in one study abstained from receiving rewards in order to be in control (Bucknoff & Metcalfe, 2020). From a social psychology perspective, these findings do seem to align with the *self-efficacy* concept by Bandura (1977) and the emphasis on self-efficacy as essential for human functioning and wellbeing.

Lately, a drift towards emphasizing the role of active involvement with engaging stimuli has been observed in research on curiosity and internally motivated behaviour (Clark, 2018). Furthermore, it has been pointed out that curiosity should not only be viewed as a perceptual phenomenon, but also a motivational feeling preparing an individual to actively engage in activities aimed at reducing the uncertainty (Van de Cruys et al., 2021). This notion about the function and influence curiosity has on perception and behaviour addresses the agentic aspects of curiosity.

By considering the evident similarity in dopaminergic activation associated with curiosity, reward and agency, as well as the inconclusive results from earlier experiments (Marvin & Shohamy, 2016), an alternative to the EIG-model, the agency-model was proposed (Metcalfe et al., 2021). It predicts that not only are humans willing to expend cognitive and physical resources to obtain non-instrumental knowledge, but participants will prefer to allocate autonomous control even when not necessary. In a study conducted by Metcalfe et al. (2021) three experiments were conducted to contrast the two models. Essential to all the experiments was the choice of either passively receiving the answer to a trivia question, or to ask for hints in the form of letters starting with the first, then the second, and so on. If participants did choose to see hints, even when the possibility of receiving the answer without working for it was also an option, this would support the agency-model. In all three experiments a preference for seeing hints were positively correlated with curiosity. Also, the duration of the time from the participants were presented with the question until they revealed the answer, was positively correlated with increased curiosity. These findings conflict with the assumption that only the rewarding properties of the answer are involved in curiosity-driven exploration and suggests that there is also an aspect of need for self-efficacy involved.

Even if curiosity incentivizes allocation of cognitive or physical resources, this does not mean that anticipation of rewarding information is not a part of the equation. If removing the rewarding stimuli, the goal-directedness of the behaviour would vanish, and so, removing the initial motivation for effort allocation. The agency-model is therefore not to be viewed as an attempt to remove reward from the equation, but simply adding an additional element to the understanding of how the decision-making process in a curious state affects behavioural choices.

Need for agentic involvement with engaging stimuli might add important nuances to our understanding of the motives underlying information-sampling when in a curious state. However, the agency perspective does not explain how intrinsic characteristics of informational input and how experiences related to that information is evaluated, eventually guiding the decision-making process involved in curiosity-driven exploratory behaviour.

1.4 Operationalization

In a seminal paper, George Loewenstein (1994) examines how curiosity has been conceptualized and researched throughout history. According to Loewenstein the trend in psychological research on curiosity was shifting away from trying to understand and explain the underlying causes of curiosity during the 1960s, to focusing on its dimensionality in the following. Up until this day, there is still no consensus concerning the underlying cause of curiosity, and researchers are still puzzled by the apparent non-instrumental nature of curiosity. Also, the dimensionality of curiosity is still a field of research. However, a relatively new line of research has developed during the last two decades or so, investigating the decision-making process of curiosity (Gottlieb & Oudeyer, 2018). In the wake of this new directional turn, challenges related to operationalization of curiosity, measuring it, and evoking it in experimental settings have proven to be a challenge. In the following section I will review how researchers previously has tackled the problem of inducing curiosity in an experimental setting. I will also review the unique challenges when creating the paradigm in this study and how they were solved.

1.4.1 Evoking curiosity

Research dealing with external reward and reinforcement mechanisms has one of the longest traditions when it comes to experimental research on behaviour and motivation (Pavlov, 1906). However, switching from external rewarding stimuli to intra-individual stimuli poses a major challenge for scientists. External rewards do all have in common that they contribute to increased chances of survival, either by providing nutrients (food or hydration) or by increasing likelihood of mating or survival through monetary values. In contrast, curiosity does not, by definition, carry any instrumental advantage. When aiming for evoking curiosity in an experimental situation, according to the principles from the information gap framework researchers ought to follow two guide lines: 1) some sort of uncertainty must occur, and 2) it is essential to make the individual for which curiosity shall be evoked aware that there are ways to diminish this uncertainty.

Visual images and written stimuli materials have been the most used methods for evoking curiosity. Building on the theoretical frameworks for curiosity, which states that certainty is a crucial prerequisite for curiosity to emerge, one way to induce curiosity has been to artificially modulate objective properties of stimulus material in such a way so that it affects experienced uncertainty.

However, monitoring certainty across different participants poses certain problems: individuals will have different initial knowledge, cognitive resources and experiences that will affect the novelty of informational inputs. One example on how uncertainty is used to create non-instrumental stimuli material is in the blurred image-paradigm by Jepma et al. (2012). There, researcher developed a stimulus set consisting of drawings of everyday objects which were modified through a blurring process leaving the original object in the image less recognizable to the participants in the study. Participants were then shown images in the following sequences: 1) blurred image, followed by the corresponding unblurred image, 2) blurred image, followed by the non-corresponding unblurred image, 3) unblurred image, followed by the corresponding blurred image, and 4) unblurred image, prior to the corresponding unblurred image. While seeing these sequences in a randomized order, brain regions associated with curiosity, such as the anterior insula and the anterior cingulate cortex (ACC), were scanned using (fMRI). The study found compelling evidence that the blurred images followed by the corresponding unblurred image elicited the most robust activation. This indirectly showed that curiosity can be induced using modified images.

Another approach has been through objective calculations of complexity, as demonstrated by Terwilliger (1963). By creating increasingly complex patterns of lines he was able to demonstrate changes in participants aha-experiences. Although no measurement of curiosity was recorded in the experiment one could speculate about a possible relation, as there is evidence for correlations between aha-experiences and curiosity(Van de Cruys et al., 2021). Another approach is seen in Van Lieshout et al. (2018), where they found a positive correlation between curiosity and uncertainty about the probability in the outcome of a lottery task. Yet another operationalization of non-instrumental information is in the form of trivia questions (Marvin & Shohamy, 2016). Even though the information in trivia questions is randomized, thus avoiding accumulation of information in specific domains, being exposed to trivia questioning in an experimental situation would still be more useful for someone who regularly attends quiz-nights.

1.4.2 Measuring curiosity

Because of the subjective nature of the curiosity experience there is no consensus in terms of how curiosity measurements should be conducted. Therefore, numerous attempts at capturing the phenomenon can be found in the literature. Self-report, brain-imaging and behavioural measures are commonly used.

Self-report has the advantages of capturing the experienced aspects of curiosity. It is also an efficient way to record curiosity in an experimental situation and is therefore a commonly used method for measuring curiosity. Self-report measures are on the other hand vulnerable to the typical fallacies of self-reports, namely extremity skewing, mid-pointing, social desirability bias (faking good), faking bad or acquiescence (passively accepting) (Furnham & Henderson, 1982). Self-report are also shown to correlate poorly with behavioural measures (Dang et al., 2020), indicating that they could fail to capture the experienced counterpart of a behavioural pattern. The information from self-report should therefore be interpreted with caution. This is not to say that self-report is meaningless, as this failure of alignment between behavioural and self-reported measures can enable broader understanding of psychological phenomena, including curiosity.

At last, the behavioural patterns, in terms of exploratory behaviour (Jepma et al., 2012), willingness to invest physical resources (Goh et al., 2021), waiting behaviour (Marvin & Shohamy, 2016) or guessing (Metcalfe et al., 2021), to name a few, have been measured as a manifestation of phenomenal curiosity. Exploratory behaviour does, however, not necessarily translate to curiosity per se. As we have seen, various theories regarding curiosity would predict distinct patterns of behaviour. For instance, the EIG-model as opposed to the agency-model, or the novelty-model compared to the prediction error-model. This diversity in theoretical perspectives can complicate our interpretation of what the recorded behavioural patterns truly reveal about curiosity.

1.4.3 Mooney images and semantic entropy

Approaching the problem of evoking and measuring curiosity form a novel perspective, researchers developed a way of calculating uncertainty for a stimulus material on a group-level (Van de Cruys et al., 2021. See material section of this article for more thorough description). Put shortly, participants were exposed to images that had been modified so that the original content beared little resemblance of its original content in a procedure called 'mooneyfication' (Mooney & Ferguson, 1951). That is, blurring the images and then thresholding them, leaving a patchwork of black and white spots. The participants were then asked to guess about the content. The guesses were gathered for each image and then, using the definition of entropy formulated by Shannon (1948), a proxy of the *semantic entropy* for each image was calculated. This calculation is formed based on the assumption that participants form hypotheses about the world around them based on the informational input received through their senses. If the informational input conveys a clear meaning, people become certain about their understanding of the information they are receiving. This certainty is formalized through a clear hypothesis about the content of the informational input. If the informational input, however, is not clear, people will not experience the same certainty about the informational input they are receiving. Thus, they will have numerous possible candidate hypotheses, equally likely to explain the informational input. Therefore, the informational gap will be proportional to the quantity of possible hypotheses about a certain informational input.

The semantic entropy estimation based on Shannons formula, utilize the variable plurality of hypotheses for different MIs, and calculates a coefficient of the semantic entropy for each MI. The plurality of hypotheses which are hypothesized to reflect the ambiguity of the informational input is captured in a coefficient. Evidence supports this assumption, showing that the participants in an experimental study were more curious about content with higher semantic entropy, as if it captured a potential information gap across participants. Based on the compelling evidence showing that the MI reliably cause curiosity, uncertainty, and aha-experiences, and as well as the objective calculation of semantic entropy, using these stimuli could be an ideal way of evoking curiosity in my experimental procedure.

1.4.4 Measuring exploration

As pointed out by Gottlieb & Oudeyer (2018), conducting experimental research on curiosity, and especially when it is related to active interrogation, diverge from how experimental paradigms typically have been developed. Where psychological research in laboratories typically have exposed test participants to a fixed stimulus eventually measuring the response of the participant to that stimulus, this setup is not appropriate when investigating how participants allocate resources. The paradigms used when conducting research on curiosity and active interrogation therefore typically diverge from the classical experimental trajectory as they increase behavioural freedom for the participants. Examples of such design is seen in paradigms by Metcalfe et al. (2021) and Petitet et al. (2021). An obvious positive aspect of increased behavioural windows is the ecological validity of the observed exploration: the participants are faced with a more natural behavioural environment than the categorical choices which traditionally is associated with experimental research. Another positive aspect is the accumulation of larger datasets enabling researchers to examine information sampling from novel perspectives. The increased amount of data points is also somewhat of a drawback, as it can complicate interpretation of the data.

Here we have shown that the burgeoning literature have provided insights into how external properties of information, as well as experiential uncertainty are related to self-reported curiosity and exploratory behaviour. We have also highlighted parts of the field being underexplained in the literature: 1) the role of agency in relation to curiosity-evoking situations, 2) the relation between certainty, curiosity, and exploration, and 3) the relation between aha-experiences and exploration.

1.5 Main hypotheses

In this study participants were exposed to a situation in which they were forced to choose between expending or preserving their cognitive resources. Their behavioural patterns were then compared with self-reported measures of curiosity, certainty, and aha-experience, as well as a crowd-sourced proxy for subjective certainty, to evaluate whether behaviour is best predicted by the EIG-model or the agency-model. Based on the evidence and theories elaborated in the previous sections, four main hypotheses divided into sub-hypotheses were formulated:

1 Certainty and curiosity

Rationale for hypothesis:

The first hypothesis is concerned with the relation between certainty and curiosity. I expect to observe a negative, linear correlation between certainty and curiosity, as seen in the study by Van de Cruys et al. (2021), which deploy the same stimuli material to evoke curiosity, as well as the same measures for certainty and curiosity.

a) Introspective curiosity will be negatively correlated with certainty.
b) Semantic entropy will be negatively correlated with curiosity.

Interpretation:

If the sub-hypotheses are supported by the results, the novelty-model would best describe the data.

2 Curiosity and exploration

Rationale for hypothesis:

The second hypothesis is concerned with the correlation between curiosity and exploration. Recent findings from curiosity research have been interpreted as evidence that curiosity motivates exploration, contrary to what previous theories have stated (Metcalfe et al., 2021).

a) Exploration will be found on some trials.
b) Exploration will be positively correlated with self-reported curiosity.
c) Introspective certainty will be negatively correlated with exploration.
d) Semantic entropy would positively be related to exploration.

Interpretation:

Rejection of the null hypothesis on the sub-hypotheses would lend support to the agency-model for curiosity.

3 Aha-experience

Rationale for hypothesis:

The third hypothesis address exploration as a contributor to the aha-experience. Aha-experiences have earlier been found to be positively related to curiosity (Van de Cruys et al., 2021). However, based on one definition of aha-experiences, stating that aha-experiences result from surprising and immediate insights, one would predict cognitive work towards a feeling of insight to modulate the aha-experience. More specific, working towards insights should increase the aha-experience.

a) The relation between curiosity and exploration is influenced by exploration.

Interpretation:
If the hypothesis is supported by the results, autonomous exploration could be interpreted to be a co-contributor to the aha-experience, along with curiosity.

4 Traits and exploration
Rationale for hypothesis:
As curiosity is related to individual traits such as intelligence, and personality (Silvia & Christensen, 2020), psychometric measures of intelligence and desire for cognitive work should be correlated with exploration.

a) NCS-scores will be positively correlated with exploration.
b) HMT-scores will be positively correlated with exploration.

Interpretation:
Rejection of the null hypothesis would indicate trait to affect exploratory curiosity-driven behaviour.

2 METHOD

2.1 Participants

32 healthy volunteers participated (20 women, 12 men, M_{age} = 26.1, SD = 5.22, age range: 21-44 years), in return for a gift card worth 200 NOK. The number of volunteers were chosen based on previous experiments revealing reliable within-person variability in their curiosity ratings with groups of between 25-30 individuals (Bloom et al., 2018). Participants were recruited through posters hung around on campus of the University of Oslo, and through the university's official website for current research projects. One participant was left out of the analyses because they chose to explore on all but one trial.

2.2 Material

2.2.1 Images

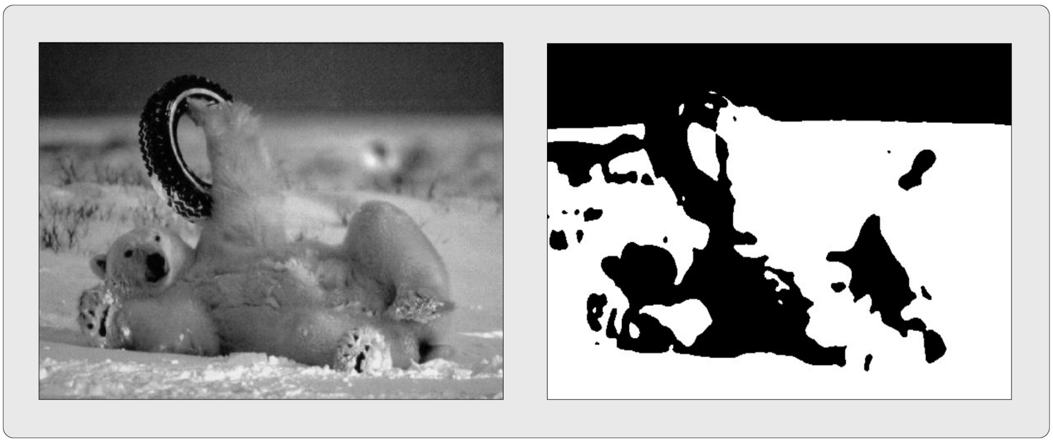
A selection of images from the Caltech 256 (Griffin et al., 2007) and the MemCat (Goetschalckx & Wagemans, 2019) image databases were further processed by Van de Cruys et al., (2021), modifying the procedure from (Imamoglu et al., 2012). The Caltech-256 database consists of 80 images for 256 everyday object categories, used in the benchmark process for object recognition in artificial intelligence. The MemCat database consists of collections of 2K images for five broad-level semantic categories (sports, food, vehicles, animals, and landscapes). The processing of the images is called "Mooneyfication", after its inventors Mooney & Ferguson (1951), and refers to the process of blurring the greyscale images, then applying a two-tone filter which converts the greyscale values above a given threshold to black, and those below to white. The Otsu method was used to determine the threshold, maximizing the variance between pixels of the two classes separated by the threshold (Walt et al., 2014).

After the Mooneyfication process, a selection process for identifying the images best suited for evoking curiosity excluded images in which the original object was still obvious after Mooneyfication, the images containing less or more than one identifiable foreground object, and the ones in which no or little discernible structure were left after the processing. The remaining images, consisting of 755 pairs of image pairs, where then evaluated by 8 participants, in a pre-trial, responding to whether they recognized the object in the Mooney image, and the strength of their aha-experience when being presented with the solution (7-point likert scale). The images which were recognized in > .80 of trials were excluded, and only the images evoking > 3.7 aha-experience were kept. For a more in depths description of the procedure, see (Van de Cruys et al., 2021).

The stimulus material in this study consisted of 203 image pairs (for an example of an image pair, see Figure 2.1). Each pair contained a grey-scale image (GSI), and a Mooney image (MI). Each image pair depicted an object or activity that could be classified into one out of six broad-level categories: animals, inanimate objects, plants, sports, vehicles, and food. As the MI of each pair were based on the GSI in that pair, each MI retained some of the features from of the GSI, although to a variable extent. The amount of information in a MI which had a predictive value when identifying the corresponding GSI, varied across the pairs. Therefore, every image pair in the dataset could be sorted on an axis, ranging from easy to difficult, in terms of how difficult it was to predict the content of the GSI based only on the information in the MI.

When an image of an object is manipulated (mooneyfied), the informational output of the image is obstructed, creating a discrepancy (an information gap) between visual knowledge of the observer, and the input it is received through the visual system. The magnitude of the discrepancy will depend on how much manipulation is done to an image. Other aspects, such as characteristics of the image before manipulation, and individual differences between participants in terms of prior knowledge and behavioural differences will affect the size of an information gap as well. This will hopefully affect the feeling of curiosity.

Figure 2. 1 Grey scale image and corresponding mooney image



Note: Figure illustrating an image pair, consisting of a GSI (on the right), and a MI. This image pair was not part of the 203 image pairs used in the experiment.

2.2.2 Semantic entropy

Another essential feature of the chosen stimulus material was that a proxy for the entropy of the pictures had been created, referred to as semantic entropy. Semantic entropy is based on the frequency and distribution of the guesses of all participants for an image calculated with Shannons definition of entropy (Shannon, 1948):

$$H(X) = -\sum P(x_i)log(P(x_i))$$

In the definition, *x* represents the probability of guess *i* to be correct. To calculate x, a fuzzy algorithm built on the Levenshtein Distance method was used to evaluate whether a guess was unique, or matched other guesses made for the same image. Each guess found to be unique, was then added as a new item to a list consisting of all unique answers for that image, setting a frequency counter for that guess to 1. If a guess did match an item on this list of guesses, the frequency counter for that particular answer was increased by one. After all guesses had been evaluated, the probability for an answer to be correct, $P(x_i)$, was calculated by dividing the frequency of one answer for an image by the total number of guesses for that image. To calculate the semantic entropy for the image, Shannons formula was applied for each unique probability, before the results from these calculations were summarized, and the number signs were reversed. This process was repeated for each image.

Uniquely for this estimation of entropy is that it is independent of the introspective evaluation of certainty, and instead relies on the knowledge participants associate with the stimuli materials. In this way, the calculations reflect the plurality of hypotheses across participants, which here is hypothesized to be a proxy of how hypothesis generation is distributed within each single participant. This also means that it is not prone to the typical sources of error associated with self-reported measures, such as initial elevation of responses (Anvari et al., 2022), motivational biases (Furnham & Henderson, 1982), as well as measuring errors caused by participants failure to translate the magnitude of information gaps into values on a Likert-scale. By both using Likert-scale measures of introspective certainty and crowdsourced estimation of entropy, a comparison between the two was enabled, which were central to the research questions in this study. Instead of calculating new values for the semantic entropy measure, we made use of calculation provided by Van de Cruys et al., (2021) since these were based on a large sample (N = 280), and also allowed us to test whether such an estimation of image properties could be transferred across participant samples.

As shown by Van de Cruys et al. (2021), this image set is well suited for the aim of this experiment as it, at least partially, solved two problems which were central to the research question. Firstly, the stimulus material evoked curiosity across individuals without the need for specific prior knowledge. Secondly, for each image pair in the stimulus material a quantitative, group-based measure of the entropy had been calculated.

To verify the effectiveness of the stimuli material to evoke curiosity, a part of the analytical procedure in the current experiment consisted of replicating central analytical procedures from Van de Cruys et al., (2021), that is, correlational analyses between the introspective variables and semantic entropy. Correlations was expected to exhibit the same directionality as observed by Van de Cruys et al., (2021). The results are shown in the results section.

2.2.3 Hints

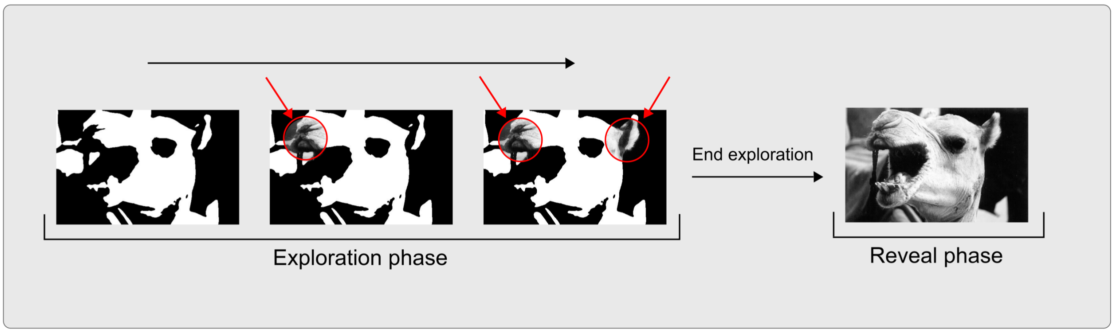
Expected reduction of uncertainty is central to our understanding of curiosity (Loewenstein, 1994; Metcalfe et al., 2020; Van de Cruys et al., 2021), and is here interpreted as a behavioural measurement for agency (Bandura, 1977; Bjork & Hommer, 2007; Bucknoff & Metcalfe, 2020; Clark, 2018; Van de Cruys et al., 2021). Implementing an option of autonomous exploration using the visual stimuli material used in this study has earlier been pointed out by Van de Cruys et al. (2021) as a possible future experiment. Here we have designed a choice between exploration and revealing the solution to the stimuli material. If choosing to explore, one gets the opportunity of freely remove the thresholding in a circular location in the picture, the size of 40 pixels, equivalent to ~3 degrees of visual angel.

The process of deciding the characteristics of a hint were guided by the aim of creating a problem which were judged as solvable by placing hints and by investing cognitive resources. Therefore, a hint had to have just the right amount of diagnosticity (reliability) (Gottlieb et al., 2020). If too much information were revealed with each hint, the exploration process could hardly be described as a problem-solving task, as the solution would be obvious to the participants after only a few hints and without notable investment of cognitive resources. If the information in a hint was too small, i.e. low diagnosticity, the problem would be unsolvable, or require more effort than participants would be willing to use. Indeed, a variability in responses was the goal, meaning that willingness to request hints should vary across trials and participants, possibly because of the semantic entropy.

To create hints with just the right amount of diagnosticity, three participants were recruited for test sessions. The test sessions were identical to a regular session, apart from the hints having slightly different characteristics for each participant, modulating the diagnosticity of the hint. The hints did vary across *size* and *opacity*: size refers to the diameter of the hint. Opacity, on the other hand, refers to how much of the thresholding filter should be removed, revealing the blurred GSI. Participant 1 was given a hint of 20 pixels, and the two others were given hints with the size of 40 pixels. The two participants receiving 40-pixel sized hints was exposed to either 0% or 50% opacity. A qualitative examination of the behaviour of the participants resulted in the 40-pixel hints with 0% opacity being chosen for the experiment. This decision was based on the observation that the participant exposed to this configuration of size and did show the most variability in responses, i.e. exploration on approximately half of the trials.

The participants could explore for the exact duration that they wanted, under the limit of 30 seconds. To end the exploration phase the participants could simply press the space button on the keyboard. There was no limit to the number of hints during the exploration phase, enabling a participants to request the desirable amount of hints. Number of hints was included as a variable in the analyses, proving further information about the characteristics of the exploratory behaviour.

Figure 2. 2 Exploration phase



Note: Illustration of the exploration and reveal phase of the experiment. Each image in the illustration represents a new state in the experiment, based on the previous choice made by the participant. In the exploration phase the participants reveal an increasingly larger area of the GSI by allocating hints (highlighted by the red arrows). Illustrated here is the progress from the state in which no hint has been made, through the first two hints. Importantly, the hints are freely placed on any location according to the participants wish. The hints in the illustration are randomly selected to give a visual representation of how hints did look in the experiment. The red circles highlighting the location of the hints is only for illustrational purposes and were not present in the actual study. Also, the number of hints were unlimited, within the duration of 30 seconds. The experiment did automatically terminate the exploration phase after 30 seconds, revealing the solution. The participant could, however, end the exploration at any time. By voluntarily ending the exploration, the participant got a second chance to make a guess about the content before the solution were revealed.

2.2.4 Psychometric measures

As a measure of trait specific factors the participants filled out the Need for Cognition Scale (Cacioppo et al., 1984), as well as the Hagen Matrices Test (HMT) (Heydasch et al., 2017).

2.2.5 Controlling for external motivation

A fixed duration for the experiment was set to one hour. This prevented participants from choosing to reveal the solutions as a way of shortening the experimental procedure. Setting a fixed time, the option of reducing the opportunity cost of being present in the experimental situation was not influencing exploratory behaviour.

2.2.6 Administrator effects

As all the sessions were administered by me, there was no need to control for potential inter-administrator effects between sessions.

2.3 Apparatus

Participants were seated in front of a 47x29.4 cm colour, flat LED monitor with a resolution of 1920x1080 pixels and 60 Hz refresh rate. Head movements were stabilized using a chin-rest that kept the eye-to-monitor distance constant at 57 cm. The experimental paradigm was coded in PsychoPy.

2.4 Statistical methods

Generalized logistic mixed models are used for analyses where the outcome variable is dichotomous, and ordinary generalized linear mixed models are used where the dependent variable is continuous. Reciever Operator Analysis (ROC) and Area Under Curve (AUC) was calculated to evaluate the goodness of fit of the logistic models (Hosmer et al., 2013). Participant and image characteristics were used as random variables.

Aggregate scores based on the introspective scores for each image were calculated for additional analyses. Aggregating the data carried two advantages: firstly, the use of linear modelling techniques such as calculating the r-squared and the correlation coefficient, were made possible after the exploration and hint variable was transformed from a dichotomous variable to a continuous one. Secondly, since exploration and hints were on continuous scales after being aggregated it enabled comparison of how the two related to both introspective measures and semantic entropy, in terms of effect size and explained variance. Pearsons r was calculated for the per-participant analysis. When investigating quadratic relationships hierarchical multiple regression analysis was used. A relation was judged to be curvilinear if the R^2 change was significant when the quadratic factor was added to the equation. Measuring the effect of exploration on subsequent aha-experience was done by applying an interaction analysis between curiosity and exploration. Output files from the experiment in PsychoPy were transformed into complete datasets using Python. All analyses and visualization were done in RStudio.

2.5 Procedure

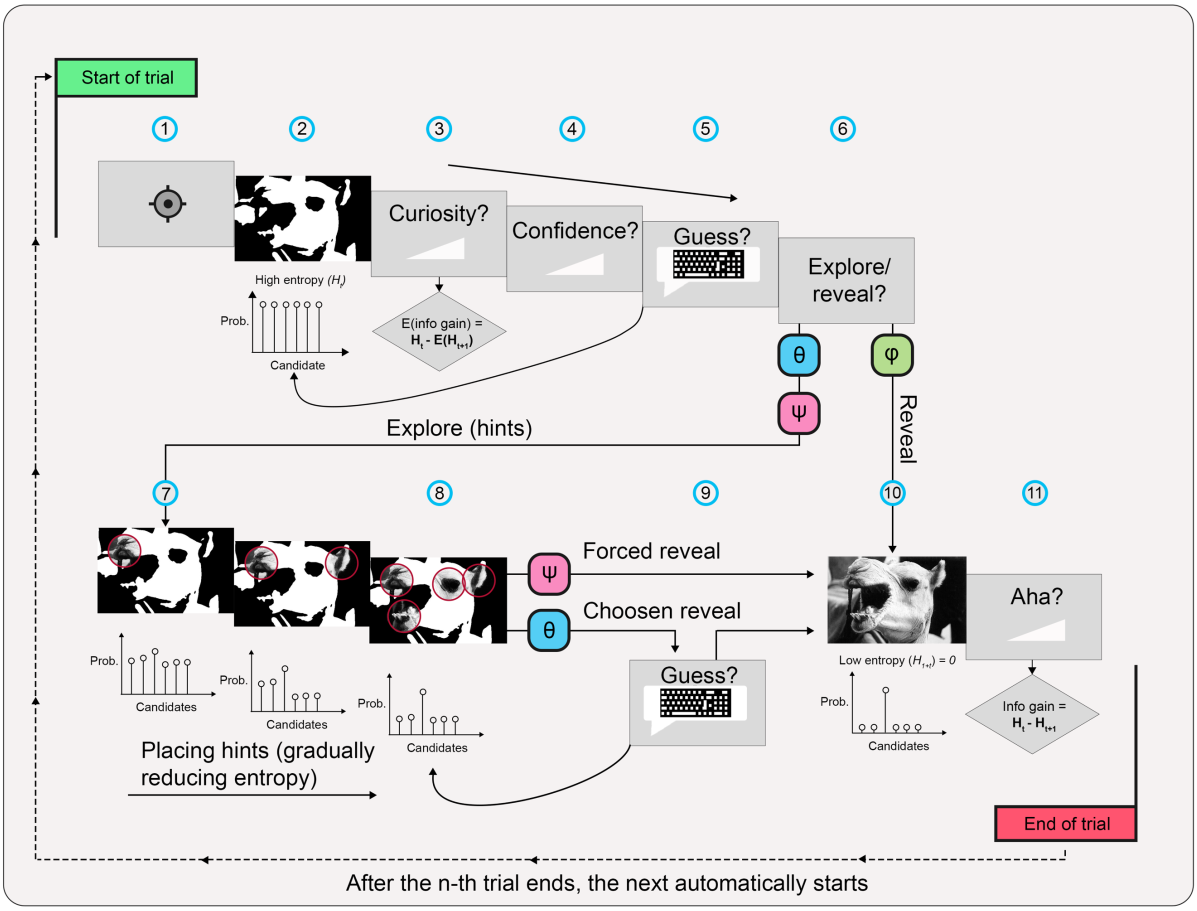
This experiment is part of a larger research project investigating decision making, memorability and corresponding physiological responses, such as eye-movements and pupil dilation. Therefore, the participant took part in this experiment as part of a session which also included various self-rapport surveys, cognitive tests, and a pupil dilation experiment, relevant for other parts of the research project. Eye-tracking for the participants were recorded while partaking the experiment. Before beginning the experiment, each participant was given written instructions about the experimental procedure.

See Figure 2.3 for illustration of the procedure. Initially in every trial a fixation cross appeared randomly in one out of eight different locations for 250 milliseconds, peripheral to where the MI would appear in the next phase for 3000 milliseconds. Immediately after the having seen the MI, the participants were asked to rank how much they would want to know the content of the picture on a 7-point rating scale (1 being "not at all", to 7 being "very much"). Then they were asked to rate how confident they were about the content of the picture on a rating scale from 1 to 7 (1 being "not at all", to 7 being "very certain"). The participants then had the opportunity to guess about the content of the picture, or to continue without guessing, after which they were given the option of either exploring the content of the picture by themselves or having the content of the picture revealed all at once. If they chose to explore the content of the picture, they selected which parts of the picture they wanted to explore, by clicking on the location with the mouse cursor. One click would reverse the thresholding for that part of the picture, revealing a circular portion of the GSI measuring 40 pixels in diameter. There was no limit as to the number of hints for an image, meaning that participants could freely request hints by clicking on different locations for the duration of 30 seconds, before an automatic reveal of the GSI. If participants wanted to quit the exploration before the exploration phase had reached automatic reveal, they could at any time press the space key to stop the exploration. This would give them a second opportunity to guess about the content. They could however also move on without guessing. After placing the guess, or choose not to guess, the GSI would be revealed. Next, regardless of their prior choice of exploring or revealing, they were asked to rate their aha-experience on a scale from 1 to 7 (1 being "absent" and 7 being "very intense"), before the next trial would automatically start.

Trial length varied as a function of participant choice to explore or reveal the pictures, but total experimental time was set equal to 60 minutes for all participants, which was therefore exposed to x number of trials. Each trail consisted of novel stimulus material which were selected through randomization. Other than that, and the location of the fixation point, each trial was identical. The participants were given instructions in three different forms before their experiment session started. First, they were shown a video presentation of the experimental procedure.

Then they were shown a flowchart of the order of the different parts of the experiment before they got to do a three test trials of the actual experiment, the same way it would be conducted (flowchart used in instructions shown in Appendix 1.1). The participants were informed about the modified pictures, that they would be asked to rate their curiosity and confidence about the pictures, and that they would have two opportunities at guessing the content in the picture: the first time after rating their curiosity and certainty, and a second time after pressing space to end exploration. In the instructions they were informed that the content in the pictures would be from six broad-level categories: animals, inanimate objects, vehicles, food, sports, and plants. Their guess, they were informed, had to be more specific than the broad-level categories. If the content in the picture was a sunflower, guessing ‘plant’ would not meet the specificity criteria. ‘Sunflower’ or ‘flower’ would, however, meet the criteria of specificity. In the instructions it was also explained that the duration of the experiment would not be affected by their choice of either exploring the images or revealing the content without exploring. They were also told that if they had to remove their face from the chinrest to move their neck at some point during the experiment, they could do that while rating curiosity and certainty.

Figure 2. 3 Illustration of the procedure



Note: Illustration showing the course of the procedure for one trial. The directions of the arrows are indicating temporal direction. *1)* Fixation point is shown initially for each trial. *2)* First exposure to the mooney image (3000 milliseconds). The plot under the image shows the theoretical probability distribution (prob.) among the candidate answers at first exposure, which translates to a high entropy (H_t). *3)* Rating of introspective evaluation of curiosity. Theoretical formula regarding the curiosity sensation is shown in the attached diamond-shape. Expected (E) informational gain is equals H_t minus the expected post-reveal entropy (H_{t+1}). *4)* Certainty rating. *5)* First guess about the content of the image. *6)* Forced choice between revealing or exploring the content. Based on this choice the subsequent course of the trial will follow either path ψ and θ or path φ. *7)* If path χ is chosen, hints can be placed freely on the image, gradually lowering the entropy. The red dots are for illustrational purposes only, to highlight the hints, and were not in the actual experiment. The plot under the three images shows how probability estimation for alternative candidates decrease with increased hints, narrowing down possible candidates. *8)* After 30 seconds of placing hints, the exploration phase is automatically terminated and trail will proceed through path ψ, where the content is revealed. If participant decides to end the exploration phase prior to the 30 second limit by pressing the space key, the experiment follows path θ, enabling the second option of placing a guess, after which the image will be revealed. *10)* The content is revealed after either of paths of one trial (ψ, θ and φ). The plot under the revealed image illustrates how entropy is finally relieved, resulting in a state where entropy equals 0 (H_{t+1}). *11)* Rating introspective aha-experience. Formula for theoretical aha-experience is shown in the diamond shaped box. Subsequent aha-experience is a sensational reflection of the actual information gain, $H_t - H_{t+1}$. The path or the stipulated line leading from “End of trial”-box to the “Start of trial”-box illustrates the circular design of the study which is fixated to one hour.

3 RESULTS

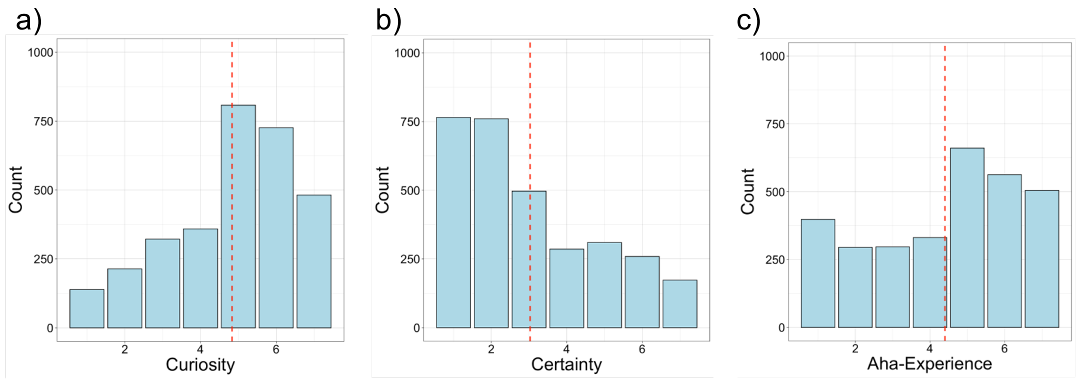
3.1 General descriptives

One participant was excluded from the analysis, as exploration was chosen on all trials but one. Among the 31 participants included a total of 3050 trials were conducted and used in the analysis. Number of trials per participant participants ranged from 66 to 125, with the mean number of trials being 97.28. The bar charts in Figure 3.1 a, b, and c reveals a skewedness in the introspective variables with both aha-experience and curiosity being right skewed, while certainty is skewed towards the left side of the chart. The means and standard deviations for variables are shown in Table 3.1.

The number of trials in which a participant chose to explore ranged from 11% to 87%, showing a relatively large difference in exploratory willingness across participants. The mean score for exploration were 58%, showing that a larger proportion of the trials resulted in exploration than merely revealing the content.

The participants made guesses in 70.84% of the first guess phase. The percentage of trials with a correct answer was 20.88% if one counts all trials in which a guess was not made as an incorrect guess. When only including trials in which a guess was made, the accuracy rose to 29.49%. In the guessing phase after exploration, participants made guesses on 47.44% of the trials, and correct guesses had increased to 34.82% if counting all trials with no guess as an incorrect answer. When counting correct guesses for only the trials in which guesses was made the percentage was 73.39. Participants repeated their first guess in the second guessing phase on 7.04% on trials. They changed their answer in in 0.34% of trials and came up with their first guess on in the second guessing phase on 17% of trials.

Figure 3. 1 Distribution of self-reported variables



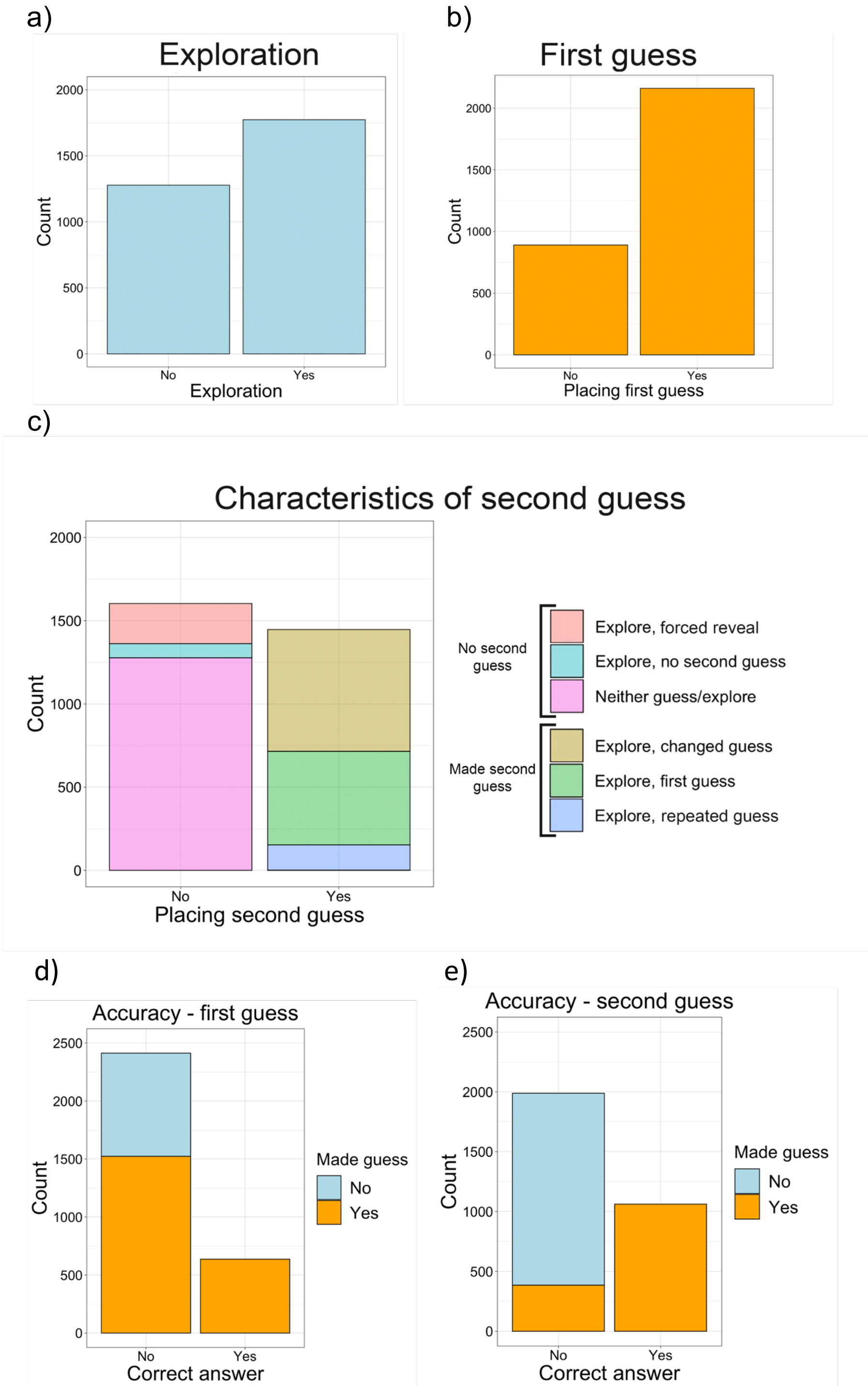
Note: Distribution of introspective variables. Self-reported score on the x-axis, and count on the y-axis. The dashed, red line shows the mean for each distribution.

Table 3. 1 Distributions for variables, including means, standard deviations and range

	Var. name	Mean	SD	Range	Var. type
Introspective self-report	Curiosity	4.83	1.66	7	Interval
	Certainty	3.03	1.85	7	Interval
	Aha	4.40	1.99	7	Interval
Self-report rating time	Curiosity	4.11	2.97	47.42	Cont.
	Certainty	2.65	2.16	33.96	Cont.
	Aha-experience	3.54	2.52	35.58	Cont.
Exploration	Exploration	0.58			Binary
	Forced reveal	0.08 (0.14)			Binary
First guess	Made guess	0.71			Binary
	Accuracy	0.21 (0.29)			Binary
Second guess	Made guess	0.47			Binary
	Accuracy	0.34 (0.79)			Binary
	Repeat	0.05 (0.07)			Binary
	Changed	0.24 (0.34)			Binary
	First	0.18			Binary

Note: means, standard deviations and range for variables. The first column indicates the category in which a variable is placed. The second column specify the variables in each category. Note that there are three variables in the second guess category, in addition to the variables "made guess" and accuracy. "Repeat" indicates that the answer in the second guessing phase match the one in the first guessing phase. "Change" means that the guess changed from the first guessing round. "First" means that no guess was made in the first round. In parentheses in the mean column, additional means based on subsets of the data is presented. For accuracy, the value in parenthesis is the mean for correct guesses when only the trials in which a guess was made is included. The number in parentheses on the row for the repeat variable is the mean when only the trials where a first guess was made was included. This is also the case for the "changed"-variable. The Standard deviation and range are only calculated for non-binary variables. The "Var. type"-column specify the three types of variables in the experiment: interval-scale, continuous scale, and binary. All binary variables are dummy coded after the same structure where 0 = "No", 1 = "Yes".

Figure 3. 2 Distributions of guesses and accuracies

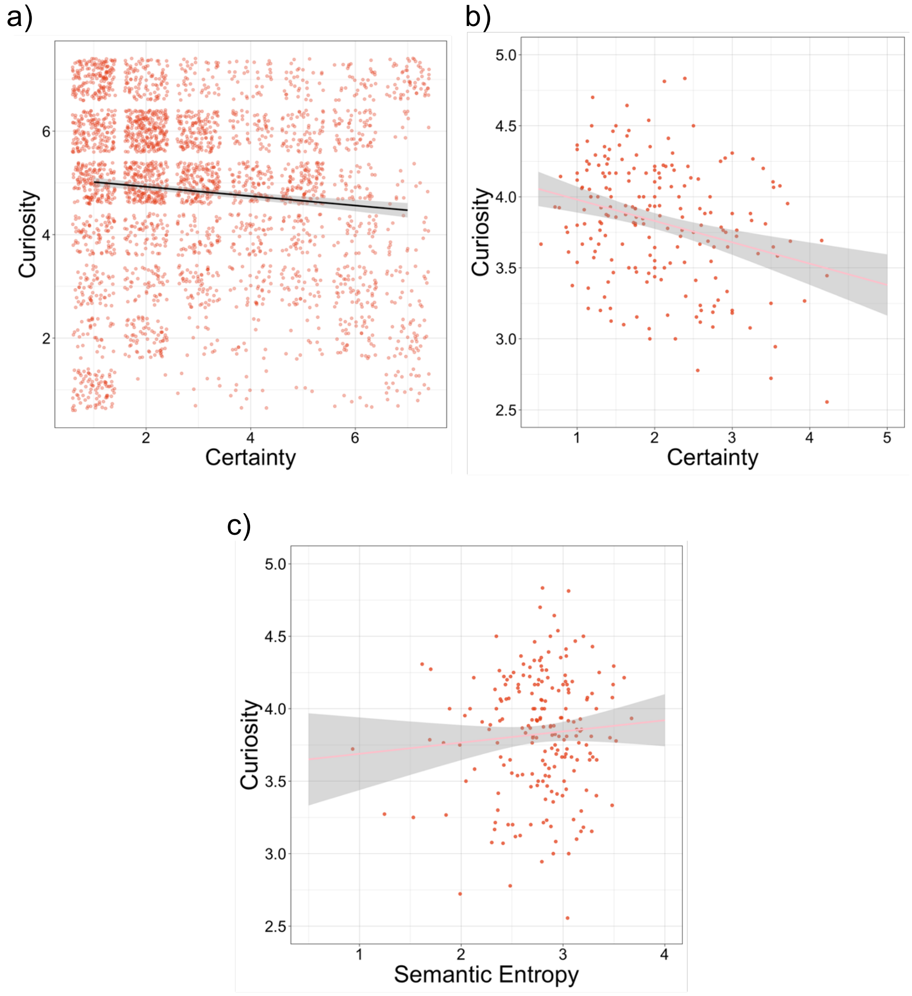


Note: Distributions of guesses and the accuracy of guesses. a) distribution of exploration. b) distribution of first guess. c) The distribution between trials for which a guess was made ("Yes") or not made ("No") in the second guessing phase is shown on the y-axis. The "Yes"- and "No"-bar is colourized to differentiate between subgroups of the trials within that bar. The pink-coloured proportion of the left bar shows the proportion where no exploration was made, resulting in the second guessing phase automatically being skipped. Turquoise on the "No"-bar indicates a choice of not guessing after having explored. Pink/red on the left bar shows the distribution of forced reveal after 30 seconds of exploration, automatically skipping the second guessing phase. For the bar on the right side of the plot the colours are differentiating between the guess being the first recorded for a trial, if it is a repetition of the first guess or if the recorded answer has changed from the first one. d) and e) shows the accuracy of the first guess and second guess, with colours indicating whether a guess was made or not.

3.2 Certainty and curiosity

The self-report scores for curiosity and certainty correlated mildly when investigating the data from each trial ($r(3048) = -.101$, $p < .001$, $R^2 = 1.0\%$). The correlation was stronger when correlating the aggregated scores for self-reported certainty and curiosity ($r(201) = -.285$, $p < .001$, $R^2 = 8.2\%$). No correlation was found between semantic entropy and curiosity ($r(201) = .078$, $p = .271$). Figure 1.3 shows plots for all models of curiosity as a function of certainty. A weak negative correlation was found between semantic entropy and certainty ($r(201) = -.152$, $p = .031$).

Figure 3. 3 Certainty and curiosity

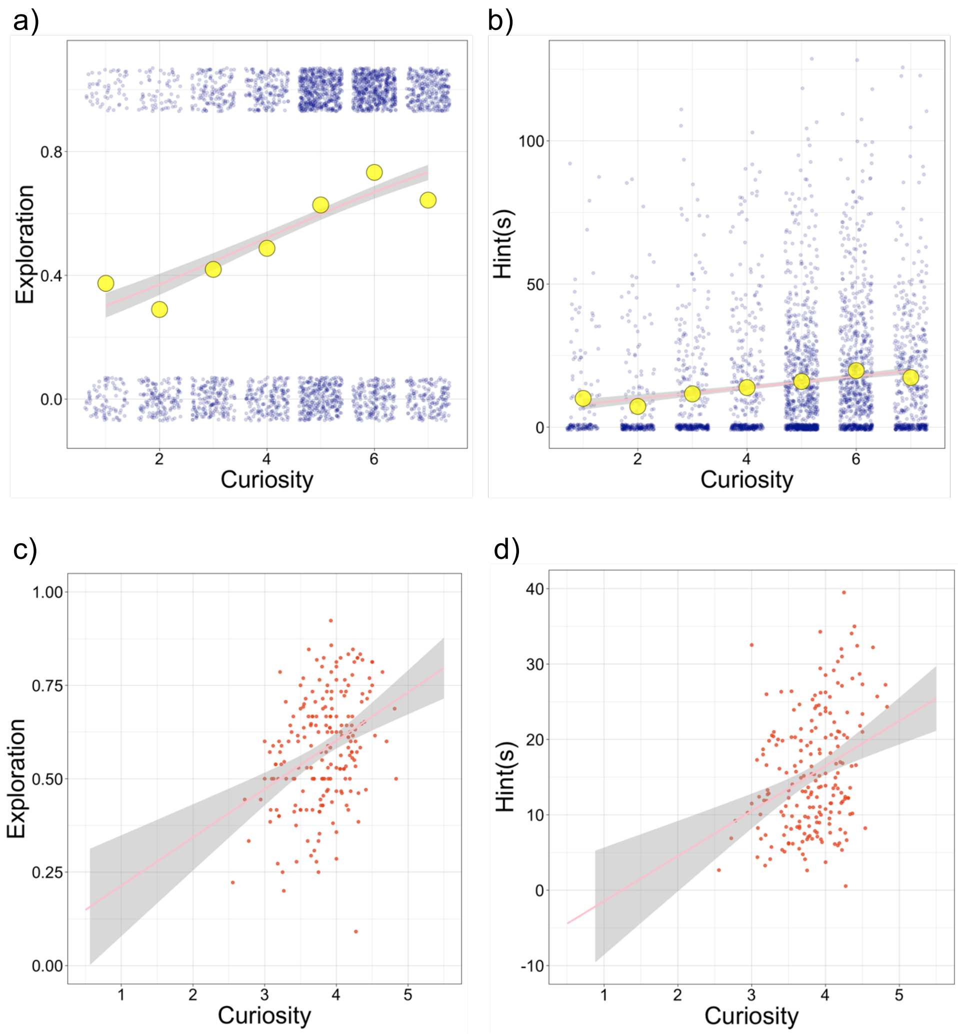


Note: Plots showing curiosity as a function of certainty for different measures of certainty. a) The red dots each represent one trial, and the black line shows the linear relationship. b) The plot shows the relationship between certainty with the aggregated data for each image. Each red dot represents one image, and the pink line shows the linear regression. c) shows the relationship between semantic entropy and curiosity. The pink line shows the relationship, and each red dot represents one image.

3.3 Curiosity and exploratory behaviour

See Figure 3.3 for illustration on the relationship between curiosity and exploration. Curiosity positively predicted exploration (b = 0.45, SE = 0.11, z = 3.94 , $p < .001$). It also positively predicted how many hints the participants requested (b = 2.05, SE = 0.64, $F(1,32) = 10.19$, $p < .001$). A ROC-curve analysis did show that the models had excellent predictive properties ($0.80 \leq AUC < 0.90$) (see Figure 3.9). The aggregated scores showed a moderate correlation between curiosity and exploratory behaviour ($r(201) = .35$, $p < .001$, $R^2 = 12.03\%$). The correlation was lower between curiosity and number of hints ($r(201) = .32$, $p < .001$, $R^2 = 10.02\%$).

Figure 3. 4 Curiosity and exploratory behaviour

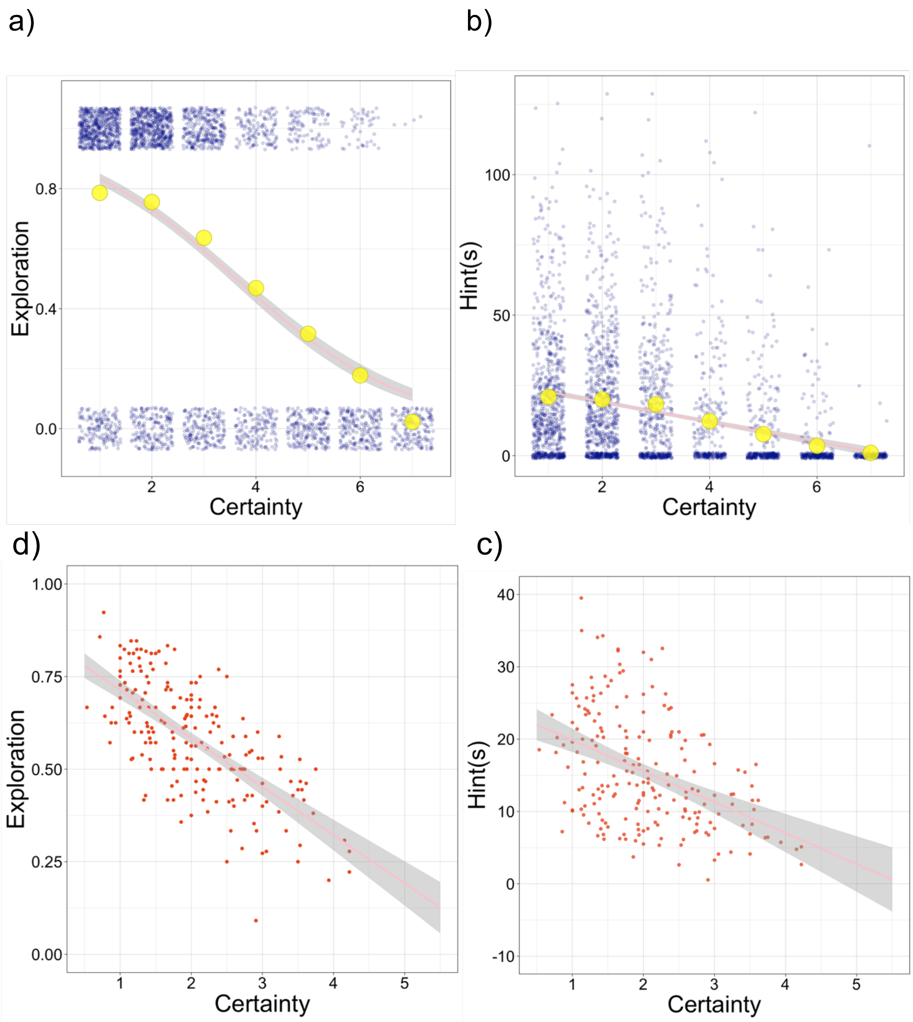


Note: a) Logistic mixed model regression showing the relationship between self-reported curiosity and willingness to exploratory behaviour. Each blue dot represents one trial. The yellow circles represent the mean for each level of exploration, and the pink line shows the sigmoid function. Figure b) shows number of hints for each trial, with the yellow circles indicating the mean, and the pink line being the linear regression. c) the aggregated scores for exploration per image. Each dot represents one image, and the line shows the best fit linear regression, and d) shows the aggregated number of hints per image.

3.4 Certainty and exploratory behaviour

There was a negative relationship between certainty and exploratory behaviour (b = -0.75, SE = 0.07, z = -10.41, p < .001). The number of hints requested was also negatively predicted by certainty (b = -3.33, SE = 0.43, F(1, 29) = 61.43, p < .001). This is illustrated in Figure 3.4. The ROC-curve analysis showed that the model had excellent predictive properties (0.80 ≤ AUC < 0.90) (see Figure 3.9 for ROC and AUC). The aggregated data showed a strong correlations between certainty and choosing to explore (r(202) = -.68, p < .001, R^2 = 46.68%). This dropped to a moderate correlation when analysing the relation between certainty and number of hints (r(202) = .43, p < .00, R^2 = 18.53%).

Figure 3. 5 Certainty and exploratory behaviour

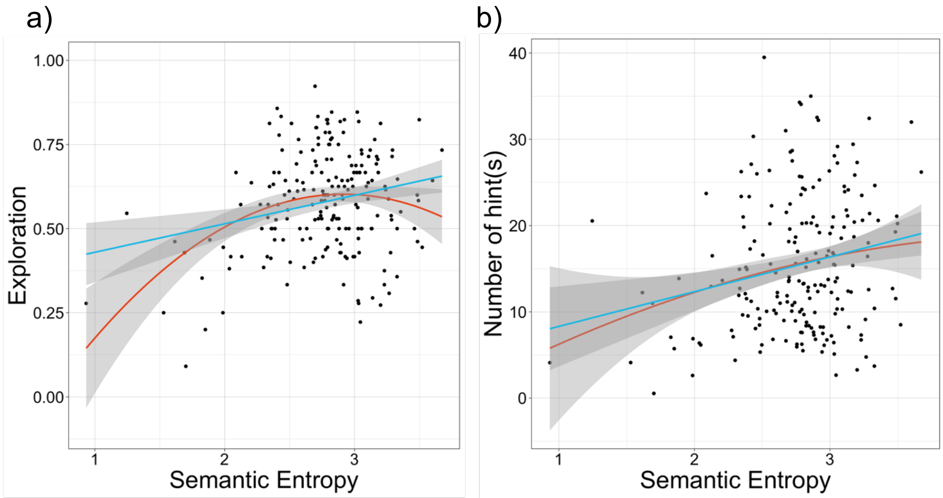


Note: a) The relationship between willingness to explore and self-reported certainty. The blue dots each represents one trial, and the yellow circles shows the mean exploration across each level. The sigmoid function is shown with the pink line. Plot b) show a scatterplot of the hints for each trial indicated by the blue dots, with the yellow circles illustrating the mean. For each level of the certainty-variable. The linear regression is shown with the pink line. c) and d) shows each images aggregated certainty score and exploration score (plot c)) and hints (plot d)). The pink line shows the linear regression.

3.5 Semantic entropy and exploration

I wanted to investigate if exploratory behaviour was linked to the estimation of semantic entropy. There was a significant correlation between semantic entropy and exploration ($r(202) = .234$, $p < .001$, $R^2 = 5.5\%$), which was also the case with number of hints ($r(202) = .214$, $p = .002$, $R^2 = 4.6\%$). When using a stepwise linear regression to add a quadratic regression, the strength of the model increased significantly for the willingness to see any hint ($r(202) = .336$, $p < .001$, $R^2 = 11.3\%$). However, there was no significant increase in explained variance when adding the quadratic regression the model for number of hints ($r(202) = .218$, $p = .586$, $R^2 = 4.7\%$).

Figure 3. 6 Semantic entropy and exploration



Note: a) scatterplot showing the correlation between semantic entropy and exploration. Each dot represents one trial, and both the quadratic and linear model is illustrated. b) scatterplot showing the correlation between semantic entropy and number of hints, with the quadratic and linear regression line in blue and red.

Table 3. 2 Semantic entropy and exploration with linear and curvilinear relation

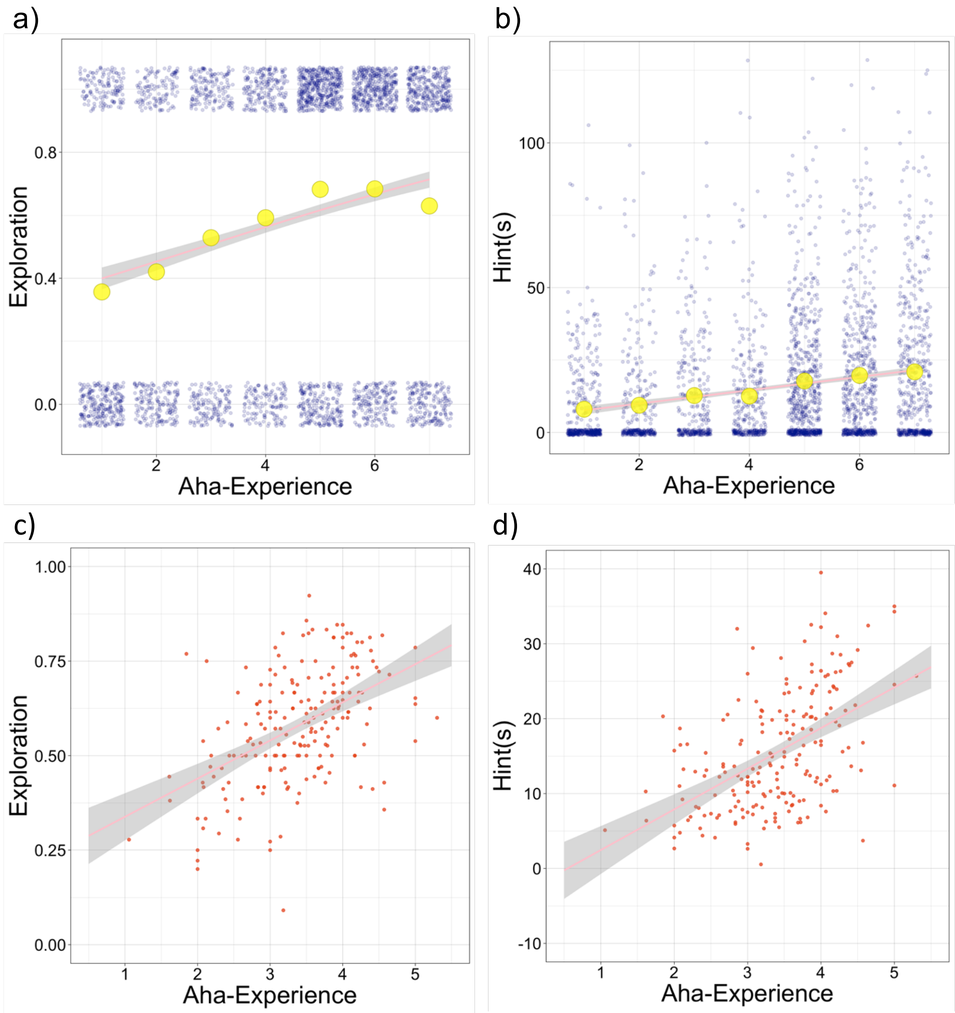
Variable	Model	R	R^2 (%)	R^2 Change	p-value
Exploration	Linear	.234	5.5	0.55	< .001
	Quadratic	.336	11.3	0.58	< .001
Hints	Linear	.214	4.6	0.46	.002
	Quadratic	.218	4.7	0.01	.586

Note: Stepwise linear modelling is used to compare the linear and quadratic models for semantic entropy and exploratory behaviour using both the explorative variable and the number of hints.

3.6 Aha and exploration

I also wanted to investigate the self-reported aha-experience and its relationship with exploration. Self-reported aha-experience positively predicted whether exploration had been chosen ($b = 0.26$, $SE = 0.05$, $z = 5.37$, $p < .001$). It also predicted the amount of requested hints positively ($b = 1.54$, $SE = 0.33$, $F(1, 25) = 21.87$, $p < .001$). A moderate correlation was found between aha-experience and exploration ($r(202) = .49$, $p < .001$, $R^2 = 24.68\%$), when using the aggregated scores for each image. A slight increase in correlation between aha-experience and exploration were seen when checking for aha and number of hints ($r(202) = .51$, $p < .001$, $R^2 = 25.92\%$).

Figure 3. 7 Aha and exploration

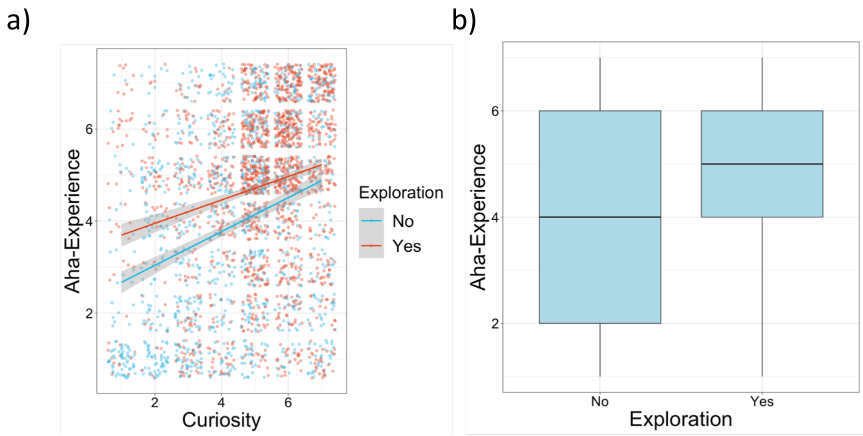


Note: a) shows a logistic regression model for the relation between introspective aha-experience and exploration, with each blue dot representing one trial. The yellow squares show the mean for the exploration variable for each level of aha-experience. Plot b) shows the number of hints for each trial indicated by the red dots, and the mean for each level indicated by the yellow circles. c) the correlation between self-reported aha-experience and aggregated exploration variables per image. Plot d) shows the same for aha-experience and number of hints.

3.7 Exploration as moderator between curiosity and aha-experience

I also wanted to see whether the seaming correlation between curiosity and aha-experience was affected by exploration. Figure 3.7a shows plots for the correlation between curiosity and aha-experiences, with exploration added as a moderator variable. A forward linear regression analysis showed that curiosity correlated mildly with aha-experience ($r(3049) = .297$, $p < .001$, $R^2 = 8.8\%$). Exploration contributed significantly to the predictive properties of the model, increasing the correlation to $r(3049) = .330$, $p < .001$, $R^2 = 10.9\%$. Adding the interaction effect between exploration and curiosity slightly increased the models predictive capacities ($r(3049) = .333$, $p = .007$, $R^2 = 11.1\%$).A Wilcoxon rank sum test was conducted to assess the differences in the aha-experience by exploration. The test did show the exploration group to have a significantly higher median ($Md = 5$, $n = 1773$) than the group trial where reveals were chosen ($Md = 4$, $n = 1277$), $U = 875513$, $z = -10.836$, $p < .001$). Illustrated in Figure 3.7b.

Figure 3. 8 Linear regression and boxplot for exploration and aha



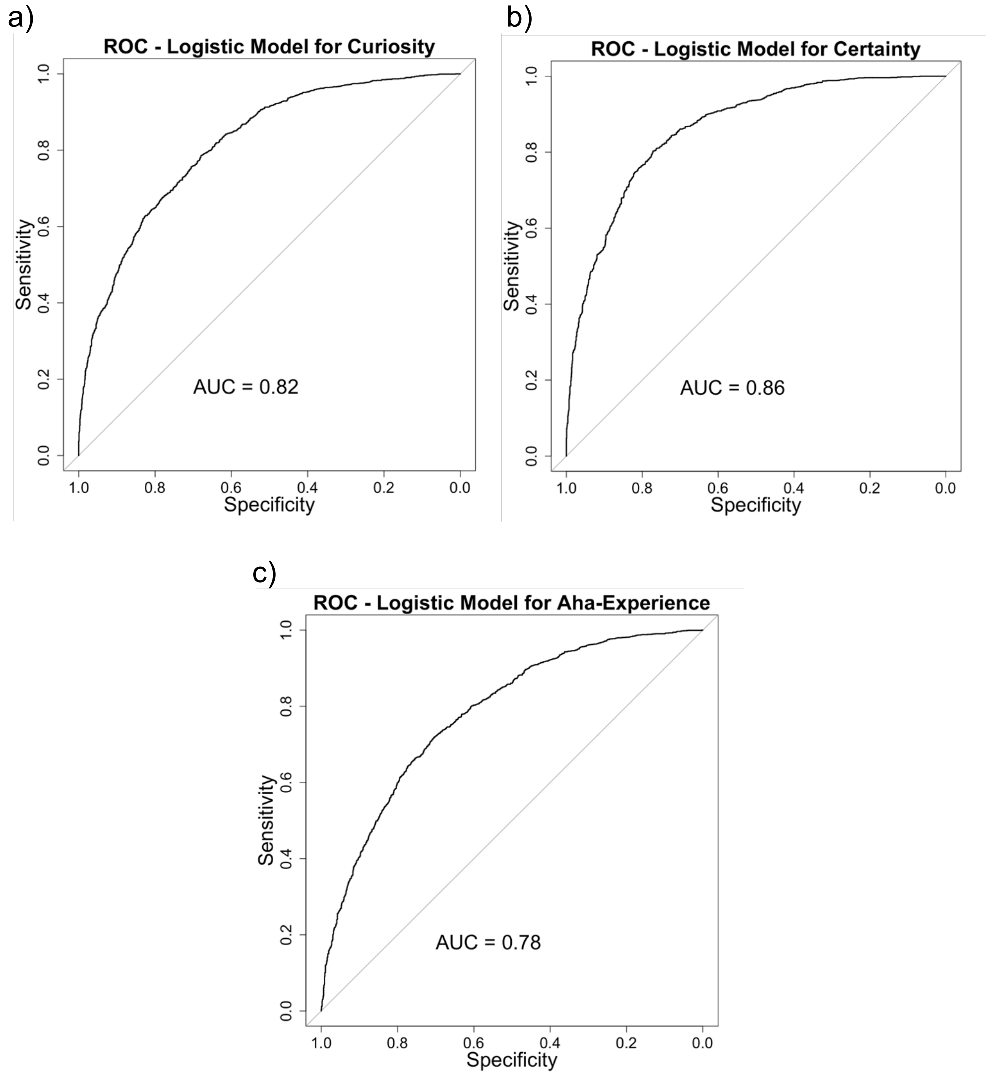
Note: a) the effect of curiosity on aha-experience, when adding exploration as a moderator variable. X-axis shows curiosity ratings and y-axis shows self-reported aha-experience. Each dot represents one trial. The colour of the dots indicate whether exploration was chosen or not. A jitter function is used to spread out the dots at each data point to better illustrate the different frequencies across the scatterplot. The turquoise line shows the relation between curiosity and aha-experience when exploration was made. b) boxplots showing the interquartile range for aha-experience for the trials where exploration was made ("Yes") and not ("No").

Table 3. 3 Interaction effects on aha-experience

	Estimation	R	R^2 (%)	R^2 Change	Standard Error	t value	p-value
Curiosity	0.36777	.297	8.8	0.088	0.02947	12.480	< .001
Exploration	1.13968	.330	10.9	0.021	0.21403	5.325	< .001
Interaction	-0.11350	.333	11.1	0.002	0.04227	-2.685	.007

Note: Forward linear regression analysis is used to add predictor values contributing to the overall predictive properties of model with a cut-off value set to $p < .05$. The model shows the interaction effect to contribute to the overall explained variance, although the r square change is small.

Figure 3. 9 ROC and AUG for logistic regressions

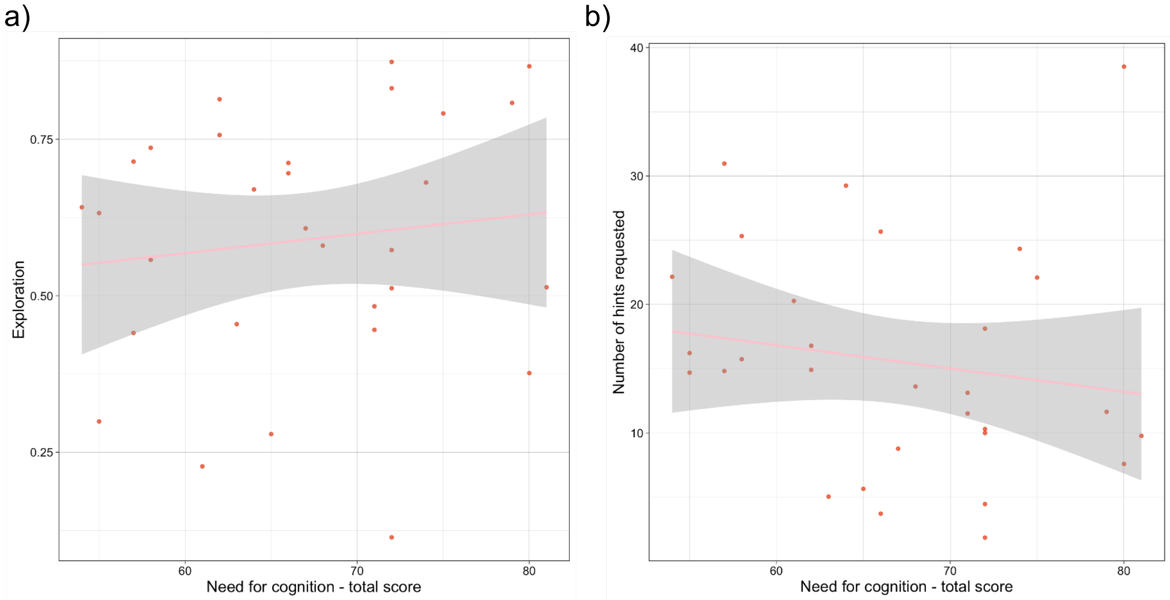


Note: ROC-curves with AUC-values for the mixed logistic regressions on introspective variables and exploration.

3.8 Per-participant analysis

An analysis of the trait measures and exploratory behaviour was also conducted. See Figure 3.10 for plots showing the relation between trait-measures and exploration. There was observed a small correlation between exploration and the NCS, although the correlation was not significant ($r(30) = .13$, $p = .51$, $R^2 = 1,59\%$). The correlation with amount of hints were negative ($r(30) = -.17$, $p = .38$, $R^2 = 2.74\%$), but not significant. Analyses revealed no relationship between HMT and exploration ($r(30) = -.02$, $p = .91$, $R^2 = 0.04\%$), or number of requested hints ($r(30) = -.01$, $p = .97$, $R^2 = 0.00\%$). Plots for the HMT-scores and exploration are not provided.

Figure 3. 10 Per-participant plots



Note: Each red dot represents one participant, and the pink line shows the linear regression model. a) shows proportion of exploration as a function of the total score of Need for Cognition Scale. b) shows the mean for amount of hints requested by a participant as a function of the total score on the Need for Cognition Scale.

4 DISCUSSION

4.1 Main findings

The findings from this study can be summarized in four points. Firstly, I found curiosity to be related to introspective certainty in a negative, linear manner, lending support to the novelty-model. Secondly, I found evidence for the agency-model, predicting that experiential curiosity is predictive of exploratory behaviour. The strongest correlation was, however, revealed to be the negative correlation between certainty and exploration. Third, self-reported aha-experience was found predicted as a positive function of curiosity. Interestingly, this relation was shown to be moderated by exploration, although the moderation effect was mild. Fourth, the stimulus material was related self-reported certainty and aha-experiences, as well as exploration. Semantic entropy did predict certainty, through a negative linear relation, and was positively predictive of aha-experiences. The relation with exploratory behaviour was best modelled with a quadratic model. However, semantic entropy did not predict curiosity.

4.1.1 Certainty and curiosity

The analyses for the correlations between certainty and curiosity are similar to the ones from Van de Cruys et al. (2021). Therefore, the results were expected to be similar to what was found in their study in terms of directionality of correlations. The results showed a correlation between certainty and curiosity which was negative, both when analysing the per-trial data and the aggregated scores. Even though self-reported certainty predicted curiosity, no correlation was found between semantic entropy and curiosity, as opposed to earlier research (Van de Cruys et al., 2021).

Although the correlations pointed in the same direction as in the original study, the correlation coefficient was much smaller. This was also to be expected, as the sample size in their experiment was almost ten times the size of the sample in this study. One study found that in replication studies, which often aim for twice the number of participants as the original study, the effect sizes on an average are half the size of that from the original study (Open Science Collaboration, 2015). I must emphasize that this study is not intended as a replication study, but research from replication studies help to contextualize the results found here. Evidence was found for a correlation between certainty and curiosity; however, certainty does only seem to explain just under one tenth of the variance in the curiosity scores when analysing the per-image aggregated scores. Aggregating scores producing stronger correlations than non-aggregated data is seen across different fields of psychology, namely psychology.

There was no indication of the quadratic relation between certainty and curiosity as proposed by the prediction error-model. As mentioned earlier, one of the principles that guides the deliberate induction of curiosity is to make it clear to participants that there is a way to obtain the information and bridge the information gap. The explanation for the quadratic relation between certainty and curiosity, as predicted by the prediction error-model, is that curiosity should occur when information sampling is evaluated to optimize predictions. In theory, high perceived certainty of information is evaluated to contribute with low potential information gain, while high certainty conveys a large information gap which is therefore judged to be unbridgeable for the individual. In designing the current experiment, a critical consideration was to create a scenario in which exploratory behaviour would not influence actual information acquisition. This allowed us to infer that if exploratory behaviour did occur, it was not serving an instrumental purpose but was inherently rewarding. This feature was essential for comparing the EIG-model and the agency-model. However, the guarantee of information acquisition, independent of effort allocation, resulted in every information gap being bridgeable. Thus, the largest information gaps carry the highest potential information gain while still being bridgeable, which could explain why a quadratic relation fail to appear.

4.1.2 Curiosity and exploration

One of the most central questions in this study was whether curiosity-driven exploratory behaviour is rewarding in and of itself, or if exploratory behaviour is merely a necessity enabling us to obtain the real reward, which is novel information? Our analyses found it to be a higher probability of seeking hints as a function of self-reported curiosity on the per-trial analysis, and this finding seems to be supported by the aggregated scores in the per-image analysis, lending support to the agency analysis.

One earlier study of waiting times and curiosity found willingness to wait to be correlated with curiosity (Marvin & Shohamy, 2016). The results were interpreted as support that effort allocation was justified because the information gain was predicted to exceed the cost. However, the study did not separate information gain and effort. To solve this Metcalfe et al., (2021) separated information gain form effort allocation, enabling investigation of whether the two contributed differently. The results showed that effort allocation was feasible, even when not necessary (described in the agency-model). This study supports the agency-model. Also, this study expands the scope of the agency-model, by showing that it does apply to visual information-sampling, as well as the verbal sampling which has previously been found.

4.1.3 Certainty and exploration

Even though curiosity was associated with increased likelihood of investing resources, the most precise predictor for exploration was certainty. There did seem to be more willingness to explore, the less confidence participants had in their interpretation of the material. Certainty has sometimes been predicted to be curvilinearly correlated with curiosity, meaning that curiosity should be highest for material of intermediate complexity, but no evidence for curvilinearity was found in this study.

4.1.4 Contribution of exploration to the aha-experience

The aha-experience was found to be predicted by curiosity. This relationship has earlier been found in at least one study (Van de Cruys et al., 2021) and was interpreted as evidence that the aha-experience is the experiential component of actual information gain.

I found the aha-experience to correlate strongly with exploration. Also, exploration did modulate the relation between curiosity and aha-experiences. This finding might seem contrary to the prevailing view that insight-learning and non-insight-learning are independent processes. While aha-experiences are linked to sudden and effortless insight learning, non-insight learning is characterised by requiring strategical and tedious work towards novel knowledge. One study has found that aha-experiences (e.g. insight learning) are unconstrained by cognitive effort (Stuyck et al., 2022), as opposed to non-insight learning.

However, exploratory behaviour does not necessarily mean that it follows the analytical non-insight trajectory as is characteristic for non-insight learning. As in the case of MIs and revealing of the GSI-content, there often is no way of acquiring the content in a gradual way. MIs often require a sudden ‘click’, when all hypotheses about the GSI-content collapse into one hypothesis, and the content is understood. Exploration could potentially, in the current experiment, be a means of narrowing the plausible candidate hypotheses, increasing the probability of experiencing an aha-experience.

Aha-experiences have also been investigated in studies of aesthetics and art, with aha-experiences being related to insight and art appreciation. At least one article has pointed out a seaming ambiguity related to art appreciation characterized by two competing needs: firstly, an observer wishes to gain insight and understanding when viewing art and secondly, the observer wants to gaining novel insights in an autonomous way (Nguyen, 2020). These needs become contradictory when gaining insight is impossible by autonomous exploration and resonation alone, requiring one to give up autonomy for example by consulting a book or an expert. On the other hand, the observers endeavour towards autonomously discovered insights will resume when the aid from supporting materials is evaluated unnecessary. In this experiment autonomously exploration carried no potential loss of informational gain. Therefore, exploration could potentially be a means of heightening the subsequent aha-experience without the risk of not obtaining information, which is often the case in natural environments.

4.1.5 Trait specific factors

A positive correlation between NCS-scores and willingness to explore was found, indicating the task presented in this experiment to be apprehended as effortful but meaningful work. Even though the correlations were non-significant, this shows that there may be some trait specific effects when it comes to willingness to explore. This could be interpreted as evidence showing that willingness to invest cognitive resources is not only situationally determined but relies on personal characteristics which makes an individual more susceptible to act in a proactive way. Even though willingness to explore is positively correlated to NCS-scores, the number of hints was correlated in the opposite direction, e.g. more hints is correlated with lower NCS-scores. A somewhat intuitive interpretation of this observation is that each hint is given more attention by participants more eager to exploit cognitive resources on each individual hint, slowing the pace of hint requests.

4.1.6 Validation of stimulus material

A correlation linking semantic entropy and exploration was found. When adding the curvilinear model to the relation between semantic entropy and exploration, the explained variance increases, meaning that a curvilinear model is better suited for explaining the relation than merely a linear relationship. This finding could lend support to the prediction error-model.

Our analyses shows that the stimuli material developed by Van de Cruys et al. (2021) reliably creates a sense of certainty, and the subsequent revealing of the solution to the MIs creates a sense of aha-experience. However, there was no significant relationship between semantic entropy and curiosity, which earlier have been reported. This raises the question of whether curiosity can be reliably predicted based on objective measures of certainty.

4.2 Further implications

Knowing that aha-experiences enhance memory, the applied consequences of this study could benefit teaching situations. Although it is difficult to evoke insight-learning, another potential way of obtaining aha-experiences is through effort allocation. This might motivate teachers to not just reveal answers, preventing students from elaborative work, as this might have negative impacts on recollecting the information.

4.3 Future studies

The following questions could be addressed in future studies.

4.3.1 What is non-instrumental information sampling?

In the burgeoning literature concerned about information sampling, the vast majority seems to draw a sharp line between the information which is sampled based on its instrumental qualities and the information sampled because of internal motivation. However, this separation is not as uncontroversial as is often implied. Is it instrumental to know all capitals in the world? Not necessarily, but one could benefit from this knowledge when deciding where to go on the next vacation, or if it comes up in a quiz. Elaborating on the notion that there might be less of a qualitative divide between instrumental and non-instrumental information sampling, one is bound to question how the two are related. One possible explanation of the link between the two is that the information which evoke curiosity does so because the information has potential relevance, hence not having a clear instrumental function. This explanation has been tested in at least one experiment (Dubey et al., 2022). In the experiment participants were presented with information of different degree of relevance and asked to rate their curiosity about the information. To manipulate “relevance” the researchers modified a piece of declarative information by adding additional information. The exemplify this in the article by manipulating information about fruit flies. In the first, non-useful, information they write that a new species of fruit flies is discovered in Malaysia, and that the discovery could lead to better insight into the locomotion in fruit flies, as it has a jumping range of 5 feet. The informational statement which are meant to have a more useful quality states that a new species of fruit flies that are found in Indonesia, and since it is discovered to share 95% of DNA with humans it could potentially facilitate a better understanding of how cancer originates in humans.

4.3.2 Curiosity as a rehearsal function for information gain

The 'potential' instrumental value of information might be the explanation for why we have curiosity. However, the exploratory behaviour associated with curiosity might be explained through another perspective. As instrumental information-sampling is bound to the acquisition of information which can be exploited for material gains, it will automatically stop when no such information is available. When no instrumental information is available, agentic exploration stops, which hinders learning about the elemental principles involved in exploration. However, curiosity could ensure that exploration endures in times when no instrumental information is available, acting as rehearsal of the exploratory behaviour function for future instrumental sampling.

4.3.3 The effect of diagnosticity on explorative behaviour

In future studies, one might want to investigate whether having a more thorough understanding of the effort required to obtain the information will affect exploration. The diagnosticity of hints might be mapped on an axis, from high on one end, to low on the other end. High diagnosticity means that one is certain that consulting hints will have valuable consequences for further knowledge acquisition (Gottlieb & Oudeyer, 2018). Low diagnosticity, on the other hand has low reliability. The predicted diagnositicy will probably also impact whether exploratory behaviour will occur.

4.4 Limitations of the study

One drawback of the present experimental design is the fact that obtaining the information is unavoidable, independent of initial curiosity or uncertainty. The choice of investing resources to obtain information is thus not a predictor of whether you will get the information or not, as would usually be the case when informational sampling takes place in natural environments. The choice regarding resource allocation will therefore not carry the same potential null outcome.

4.5 Conclusion

Going back to the initial questions: does curiosity motivate us to autonomous exploration? Based on results from our study it is tempting to answer this with “yes”. Curiosity and certainty are both stable predictors for exploration. So is the objective measure of curiosity. The stimuli material used here does cause explorative behaviour and aha-experiences but fail at evoking curiosity.

5 REFERENCES

Anvari, F., Efendić, E., Olsen, J., Arslan, R. C., Elson, M., & Schneider, I. K. (2022). Bias in Self-Reports: An Initial Elevation Phenomenon. *Social Psychological and Personality Science*, 19485506221129160. https://doi.org/10.1177/19485506221129160

Bandura, A. (1977). *Social Learning Theory*. Prentice Hall.

Baranes, A., Oudeyer, P.-Y., & Gottlieb, J. (2015). Eye movements reveal epistemic curiosity in human observers. *Vision Research*, *117*, 81–90. https://doi.org/10.1016/j.visres.2015.10.009

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain Dopamine Neurons Encode a Quantitative Reward Prediction Error Signal. *Neuron*, *47*(1), 129–141. https://doi.org/10.1016/j.neuron.2005.05.020

Berlyne, D. E. (1954). A Theory of Human Curiosity. *British Journal of Psychology. General Section*, *45*(3), 180–191. https://doi.org/10.1111/j.2044-8295.1954.tb01243.x

Berridge, K. C., & Kringelbach, M. L. (2015). Pleasure Systems in the Brain. *Neuron*, *86*(3), 646–664. https://doi.org/10.1016/j.neuron.2015.02.018

Bjork, J. M., & Hommer, D. W. (2007). Anticipating instrumentally obtained and passively-received rewards: A factorial fMRI investigation. *Behavioural Brain Research*, *177*(1), 165–170. https://doi.org/10.1016/j.bbr.2006.10.034

Bloom, P. A., Friedman, D., Xu, J., Vuorre, M., & Metcalfe, J. (2018). Tip-of-the-tongue states predict enhanced feedback processing and subsequent memory. *Consciousness and Cognition*, *63*, 206–217. https://doi.org/10.1016/j.concog.2018.05.010

Brehm, J. W., & Self, E. A. (1989). The Intensity of Motivation. *Annual Review of Psychology*, *40*(1), 109–131. https://doi.org/10.1146/annurev.ps.40.020189.000545

Brennen, T., Vikan, A., & Dybdahl, R. (2007). Are tip-of-the-tongue states universal? Evidence from the speakers of an unwritten language. *Memory*, *15*(2), 167–176. https://doi.org/10.1080/09658210601164743

Bromberg-Martin, E. S., & Hikosaka, O. (2009). Midbrain Dopamine Neurons Signal Preference for Advance Information about Upcoming Rewards. *Neuron*, *63*(1), 119–126. https://doi.org/10.1016/j.neuron.2009.06.009

Bucknoff, Z. J., & Metcalfe, J. (2020). Memory Under the SEA (Subjective Experience of Agency). In *Memory Quirks*. Routledge.

Cacioppo, J., Petty, R., & Kao, C. (1984). The efficient assessment of NFC. *Journal of Personality Assessment*, *48*, 306–307. https://doi.org/10.1207/s15327752jpa4803_13

Clark, A. (2018). A nice surprise? Predictive processing and the active pursuit of novelty. *Phenomenology and the Cognitive Sciences*, *17*(3), 521–534. https://doi.org/10.1007/s11097-017-9525-z

Dang, J., King, K. M., & Inzlicht, M. (2020). Why Are Self-Report and Behavioral Measures Weakly Correlated? *Trends in Cognitive Sciences*, *24*(4), 267–269. https://doi.org/10.1016/j.tics.2020.01.007

Daw, N. D., & Doya, K. (2006). The computational neurobiology of learning and reward. *Current Opinion in Neurobiology*, *16*(2), 199–204. https://doi.org/10.1016/j.conb.2006.03.006

Day, H. I. (Ed.). (1981). *Advances in Intrinsic Motivation and Aesthetics*. Springer US. https://doi.org/10.1007/978-1-4613-3195-7

Dubey, R., & Griffiths, T. L. (2020). Understanding exploration in humans and machines by formalizing the function of curiosity. *Current Opinion in Behavioral Sciences*, *35*, 118–124. https://doi.org/10.1016/j.cobeha.2020.07.008

Dubey, R., Griffiths, T. L., & Lombrozo, T. (2022). If it's important, then I'm curious: Increasing perceived usefulness stimulates curiosity. *Cognition*, *226*, 105193. https://doi.org/10.1016/j.cognition.2022.105193

Furnham, A., & Henderson, M. (1982). The good, the bad and the mad: Response bias in self-report measures. *Personality and Individual Differences*, *3*(3), 311–320. https://doi.org/10.1016/0191-8869(82)90051-4

Gallagher, M. W., & Lopez, S. J. (2007). Curiosity and well-being. *The Journal of Positive Psychology*, *2*(4), 236–248. https://doi.org/10.1080/17439760701552345

Gerken, L., Balcomb, F. K., & Minton, J. L. (2011). Infants avoid 'labouring in vain' by attending more to learnable than unlearnable linguistic patterns. *Developmental Science*, *14*, 972–979. https://doi.org/10.1111/j.1467-7687.2011.01046.x

Goetschalckx, L., & Wagemans, J. (2019). MemCat: A new category-based image set quantified on memorability. *PeerJ*, *7*, e8169. https://doi.org/10.7717/peerj.8169

Goh, A. X.-A., Bennett, D., Bode, S., & Chong, T. T.-J. (2021). Neurocomputational mechanisms underlying the subjective value of information. *Communications Biology*, *4*(1), Article 1. https://doi.org/10.1038/s42003-021-02850-3

Gottlieb, J., Cohanpour, M., Li, Y., Singletary, N., & Zabeh, E. (2020). Curiosity, information demand and attentional priority. *Current Opinion in Behavioral Sciences*, *35*, 83–91. https://doi.org/10.1016/j.cobeha.2020.07.016

Gottlieb, J., & Oudeyer, P.-Y. (2018). Towards a neuroscience of active sampling and curiosity. *Nature Reviews Neuroscience*, *19*(12), Article 12. https://doi.org/10.1038/s41583-018-0078-0

Griffin, G., Holub, A., & Perona, P. (2007). Caltech-256 Object Category Dataset. *CalTech Report*.

Heydasch, T., Haubrich, J., & Renner, K.-H. (2017). The Short Version of the Hagen Matrices Test (HMT-S): 6-Item Induction Intelligence Test. *methods*, *data*, 26 Pages. https://doi.org/10.12758/MDA.2013.011

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression* (3rd ed.). John Wiley & Sons, Inc. DOI:10.1002/9781118548387

Imamoglu, F., Kahnt, T., Koch, C., & Haynes, J.-D. (2012). Changes in functional connectivity support conscious object recognition. *NeuroImage*, *63*(4), 1909–1917. https://doi.org/10.1016/j.neuroimage.2012.07.056

James, W. (2007). *The Principles of Psychology*. Cosimo, Inc.

Jepma, M., Verdonschot, R., van Steenbergen, H., Rombouts, S., & Nieuwenhuis, S. (2012). Neural mechanisms underlying the induction and relief of perceptual curiosity. *Frontiers in Behavioral Neuroscience*, *6*. https://www.frontiersin.org/articles/10.3389/fnbeh.2012.00005

Kahneman, D. (2011). *Thinking, fast and slow* (1st ed). Farrar, Straus and Giroux.

Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T., & Camerer, C. F. (2009). *The Wick in the Candle of Learning: Epistemic Curiosity Activates Reward Circuitry and Enhances Memory*. https://journals-sagepub-com.ezproxy.uio.no/doi/10.1111/j.1467-9280.2009.02402.x

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2014). The Goldilocks effect in infant auditory attention. *Child Development*, *85*(5), 1795–1804. https://doi.org/10.1111/cdev.12263

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, *100*(4), 609–639. https://doi.org/10.1037/0033-295x.100.4.609

Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *35*(4), 989–998. https://doi.org/10.1037/a0015729

Kounios, J., & Beeman, M. (2014). The cognitive neuroscience of insight. *Annual Review of Psychology*, *65*, 71–93. https://doi.org/10.1146/annurev-psych-010213-115154

Lau, J. K. L., Ozono, H., Kuratomi, K., Komiya, A., & Murayama, K. (2020). Shared striatal activity in decisions to satisfy curiosity and hunger at the risk of electric shocks. *Nature Human Behaviour*, *4*(5), Article 5. https://doi.org/10.1038/s41562-020-0848-3

Leotti, L. A., & Delgado, M. R. (2011). The Inherent Reward of Choice. *Psychological Science*, *22*(10), 1310–1318. https://doi.org/10.1177/0956797611417005

Leotti, L. A., & Delgado, M. R. (2014). The value of exercising control over monetary gains and losses. *Psychological Science*, *25*(2), 596–604. https://doi.org/10.1177/0956797613514589

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, *116*(1), 75. https://doi.org/10.1037/0033-2909.116.1.75

Markant, D. B., & Gureckis, T. M. (2014). Is it better to select or to receive? Learning via active and passive hypothesis testing. *Journal of Experimental Psychology: General*, *143*, 94–122. https://doi.org/10.1037/a0032108

Marvin, C. B., & Shohamy, D. (2016). Curiosity and reward: Valence predicts choice and information prediction errors enhance learning. *Journal of Experimental Psychology: General*, *145*(3), 266. https://doi.org/10.1037/xge0000140

Metcalfe, J., Kennedy-Pyers, T., & Vuorre, M. (2021). Curiosity and the desire for agency: Wait, wait … don’t tell me! *Cognitive Research: Principles and Implications*, *6*(1), 69. https://doi.org/10.1186/s41235-021-00330-0

Metcalfe, J., Schwartz, B. L., & Eich, T. S. (2020). Epistemic curiosity and the region of proximal learning. *Current Opinion in Behavioral Sciences*, *35*, 40–47. https://doi.org/10.1016/j.cobeha.2020.06.007

Mooney, C. M., & Ferguson, G. A. (1951). A new closure test. *Canadian Journal of Psychology*, *5*(3), 129. https://doi.org/10.1037/h0083540

Murayama, K., Matsumoto, M., Izuma, K., Sugiura, A., Ryan, R. M., Deci, E. L., & Matsumoto, K. (2015). How self-determined choice facilitates performance: A key role of the ventromedial prefrontal cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, *25*(5), 1241–1251. https://doi.org/10.1093/cercor/bht317

Muth, C., & Carbon, C.-C. (2013). The aesthetic aha: On the pleasure of having insights into Gestalt. *Acta Psychologica*, *144*(1), 25–30. https://doi.org/10.1016/j.actpsy.2013.05.001

Nguyen, C. T. (2020). Autonomy and Aesthetic Engagement. *Mind*, *129*(516), 1127–1156. https://doi.org/10.1093/mind/fzz054

O'Doherty, J. P., Deichmann, R., Critchley, H. D., & Dolan, R. J. (2002). Neural responses during anticipation of a primary taste reward. *Neuron*, *33*(5), 815–826. https://doi.org/10.1016/s0896-6273(02)00603-7

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. https://doi.org/10.1126/science.aac4716

Pavlov, I. P. (1906). The Scientific Investigation of the Psychical Faculties or Processes in the Higher Animals. *Science*, *24*(620), 613–619. https://doi.org/10.1126/science.24.620.613

Petitet, P., Attaallah, B., Manohar, S. G., & Husain, M. (2021). The computational cost of active information sampling before decision-making under uncertainty. *Nature Human Behaviour*, *5*(7), Article 7. https://doi.org/10.1038/s41562-021-01116-6

Piaget, J. (1950). *The psychology of intelligence*. Routledge & Kegan Paul.

Ryan, M. P., Petty, C. R., & Wenzlaff, R. M. (1982). Motivated remembering efforts during tip-of-the-tongue states. *Acta Psychologica*, *51*(2), 137–147. https://doi.org/10.1016/0001-6918(82)90058-0

Schwartz, B. L. (2006). Tip-of-the-tongue states as metacognition. *Metacognition and Learning*, *1*(2), 149–158. https://doi.org/10.1007/s11409-006-9583-z

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Sharot, T., & Sunstein, C. R. (2020). How people decide what they want to know. *Nature Human Behaviour*, *4*(1), Article 1. https://doi.org/10.1038/s41562-019-0793-1

Sharot, T., Velasquez, C. M., & Dolan, R. J. (2010). Do decisions shape preference? Evidence from blind choice. *Psychological Science*, *21*, 1231–1235. https://doi.org/10.1177/0956797610379235

Silvestrini, N., Musslick, S., Berry, A. S., & Vassena, E. (2022). An integrative effort: Bridging motivational intensity theory and recent neurocomputational and neuronal models of effort and control allocation. *Psychological Review*. https://doi.org/10.1037/rev0000372

Silvia, P. J., & Christensen, A. P. (2020). Looking up at the curious personality: Individual differences in curiosity and openness to experience. *Current Opinion in Behavioral Sciences*, *35*, 1–6. https://doi.org/10.1016/j.cobeha.2020.05.013

Son, L. K., & Metcalfe, J. (2000). Metacognitive and Control Strategies in Study-Time Allocation. *Journal of Experimental Psychology: Learning Memory and Cognition*, *26*(1), 204–221. Scopus. https://doi.org/10.1037/0278-7393.26.1.204

Steels, L. (2004). The Autotelic Principle. In F. Iida, R. Pfeifer, L. Steels, & Y. Kuniyoshi (Eds.), *Embodied Artificial Intelligence: International Seminar, Dagstuhl Castle,* *Germany, July 7-11, 2003. Revised Papers* (pp. 231–242). Springer. https://doi.org/10.1007/978-3-540-27833-7_17

Stuyck, H., Cleeremans, A., & Van den Bussche, E. (2022). Aha! under pressure: The Aha! experience is not constrained by cognitive load. *Cognition*, *219*, 104946. https://doi.org/10.1016/j.cognition.2021.104946

Ten, A., Kaushik, P., Oudeyer, P.-Y., & Gottlieb, J. (2021). Humans monitor learning progress in curiosity-driven exploration. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-26196-w

Terwilliger, R. F. (1963). Pattern Complexity and Affective Arousal. *Perceptual and Motor Skills*, *17*(2), 387–395. https://doi.org/10.2466/pms.1963.17.2.387

Thiede, K. W., & Dunlosky, J. (1999). Toward a general model of self-regulated study: An analysis of selection of items for study and self-paced study time. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1024–1037. https://doi.org/10.1037/0278-7393.25.4.1024

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*, 189–208. https://doi.org/10.1037/h0061626

Tricomi, E. M., Delgado, M. R., & Fiez, J. A. (2004). Modulation of caudate activity by action contingency. *Neuron*, *41*(2), 281–292. https://doi.org/10.1016/s0896-6273(03)00848-1

Van de Cruys, S. (2017). Affective Value in the Predictive Mind. In T. K. Metzinger & W. Wiese (Eds.), *PPP - Philosophy and Predictive Processing*. Philosophy and Predictive Processing. Frankfurt am Main: MIND Group. https://doi.org/10.15502/9783958573253

Van de Cruys, S., Damiano, C., Boddez, Y., Król, M., Goetschalckx, L., & Wagemans, J. (2021). Visual affects: Linking curiosity, Aha-Erlebnis, and memory through information gain. *Cognition*, *212*, 104698. https://doi.org/10.1016/j.cognition.2021.104698

Van Lieshout, L. L. F., Vandenbroucke, A. R. E., Müller, N. C. J., Cools, R., & De Lange, F. P. (2018). Induction and Relief of Curiosity Elicit Parietal and Frontal Activity. *The Journal of Neuroscience*, *38*(10), 2579–2588. https://doi.org/10.1523/JNEUROSCI.2816-17.2018

Vygotsky, L. (1962). *Thought and language* (pp. xxi, 168). MIT Press. https://doi.org/10.1037/11193-000

Wade, S., & Kidd, C. (2019). The role of prior knowledge and curiosity in learning. *Psychonomic Bulletin & Review*, *26*(4), 1377–1387. https://doi.org/10.3758/s13423-019-01598-6

Walt, S. van der, Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., & Yu, T. (2014). scikit-image: Image processing in Python. *PeerJ*, *2*, e453. https://doi.org/10.7717/peerj.453

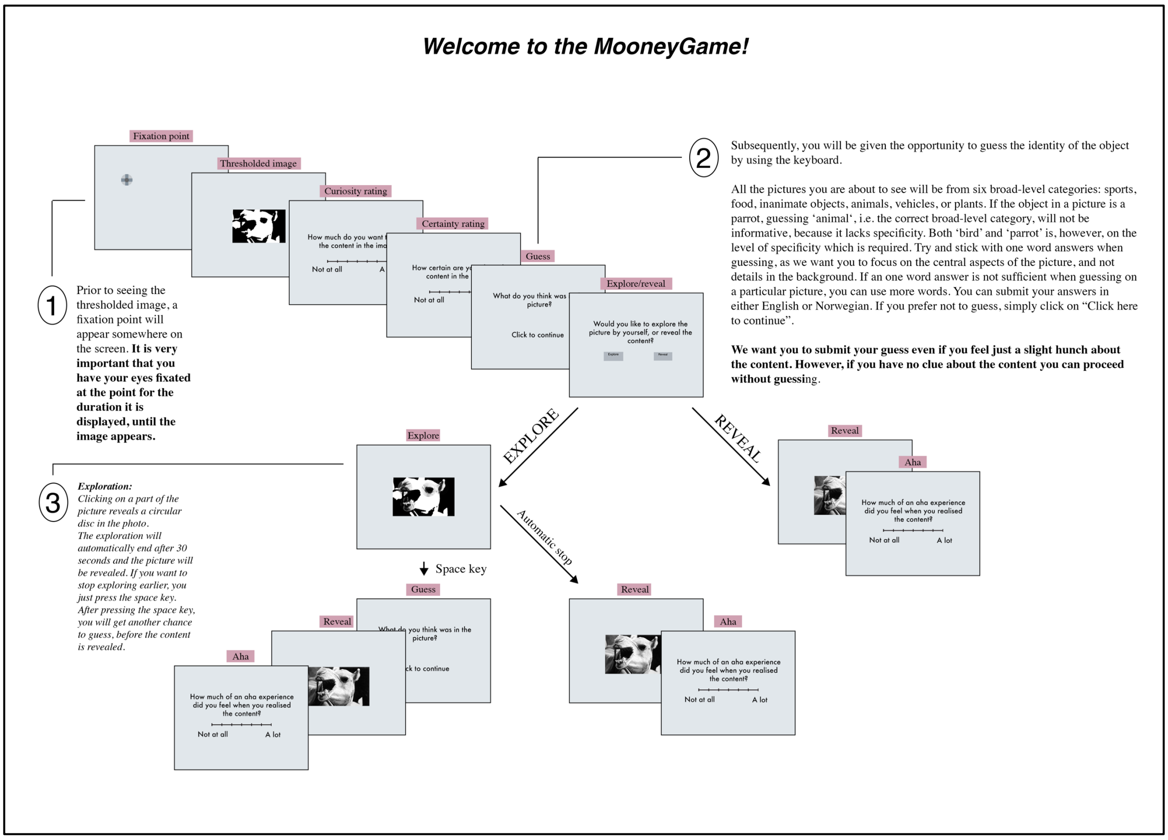
Wang, M. Z., & Hayden, B. Y. (2021). Latent learning, cognitive maps, and curiosity. *Current Opinion in Behavioral Sciences*, *38*, 1–7. https://doi.org/10.1016/j.cobeha.2020.06.003

Witek, M. A. G., Matthews, T., Bodak, R., Blausz, M. W., Penhune, V., & Vuust, P. (2023). Musicians and non-musicians show different preference profiles for single chords of varying harmonic complexity. *PLOS ONE*, *18*(2), e0281057. https://doi.org/10.1371/journal.pone.0281057

Xu, J., & Metcalfe, J. (2016). Studying in the region of proximal learning reduces mind wandering. *Memory & Cognition*, *44*(5), 681–695. https://doi.org/10.3758/s13421-016-0589-8

6 Appendix

Appendix 1. 1 Instructions - flowchart



Note: The instructions given to the participants as part of the experimental procedure.