CURIOSITY AND MOTIVATION TOWARD VISUAL INFORMATION

Erik Lundgren
I want to thank my supervisor Linus Holm for providing tireless feedback and supervision when I kept visiting the outer boundaries of the zone of proximal development.

This thesis was supported by the project “Measuring the costs and benefits of curiosity: Exploratory behavior and learning outcomes (MMW 2013.0196)”, led by Linus Holm, PhD, funded by the Marianne and Marcus Wallenberg foundation.
Abstract

Curiosity is thought to be an intrinsically motivated driving force for seeking information. Thus, the opportunity for an information gain (IG) should instil curiosity in humans and result in information gathering actions. To investigate if, and how, information acts as an intrinsic reward, a search task was set in a context of blurred background images which could be revealed by iterative clicking. The search task was designed such that it prevented efficient IG about the underlying images. Participants therefore had to trade between clicking regions with high search target probability or high expected image content information. Image content IG was established from “information-maps” based on participants exploration with the intention of understanding (1) the main theme of the image and (2) how interesting the image might appear to others. Note that IG is in this thesis not identical with the information theoretic concept of information gain, the quantities are however probably related. It was hypothesised that participants would be distracted by visually informative regions and that images independently rated as more interesting would yield higher image based IG. It was also hypothesised that image based IG would increase as a function of time. Results show that participants sometimes explored images driven by curiosity, and that there was considerable individual variation in which images participants were curious about. Independent interest ratings did not account for image based IG. The level of IG increased over trials, interestingly without affecting participants’ performance on the visual search task designed to prevent IG. Results support that IG is rewarding as participants learned to optimize IG over trials without compromising performance on the extrinsically motivated search; managing to both keep the cake and eat it.

Keywords: curiosity, exploratory behavior, interest, motivation, reward, visual search.

Nyfikenhet är en inneboende drivkraft mot belönande informationsvinster (IV). Möjligheten till IV bör kunna locka fram nyfikenhet vilket visar sig i observerbara beteenden som resulterar i IV. För att undersöka om, och hur, IV fungerar som en inneboende drivkraft, så utfördes en sökuppgift i en miljö av suddiga bakgrundsbilder vilka kunde göras tydliga genom iterativa klickningar. Sökuppgiften var designad så att den hindrade effektiv IV. IV mättes genom att skapa informationskartor över bilderna som var baserade på utforsknande av bilderna med intentionen att (1) förstå huvudmotivet, samt (2) förstå hur intressant bilden skulle vara för andra observatörer. Notera att det inte var strikt informationsteoretisk IV som mättes, kvantiteterna bör dock vara nära relaterade. Hypoteser var att deltagare skulle distraheras mot att utforska informativa bildområden, att bilder oberoende skattade som mer intressanta skulle resultera i högre IV, samt att IV skulle öka som en funktion av tid. Resultat visade att deltagare ibland styrdes mot utforsknande beteenden som resulterade i ökad IV, men att det är stora individuella variationer gällande vilka bilder som deltagare var nyfikna på. Bilder med högre intresseeskattning ledde ej till ökad IV. IV ökade över tid, intressant nog utan att påverka prestationen på den visuella sökuppgiften, vilken skulle hindra IV. Resultaten stödjer teorin att IV är inneboende belönande eftersom deltagare lyckades optimera långsiktig informationsvinst utan att kompromissa på prestation i den yttre motiverade sökuppgiften; deltagarna lyckades äta kakan och samtidigt behålla den.

 Nyckelord: nyfikenhet, utforsknande beteende, intresse, motivation, belönning, visuell sökning.
Curiosity and motivation toward visual information

Curiosity is a fundamental intrinsic motivational force towards behaviours of exploration and information seeking. Despite curiosity being important in facilitating learning (ref) and prevalent in everyday daily life, curiosity has not received much attention in psychological research (Loewenstein, 1994), and researchers in the field regard it as not well understood (Gottlieb, Lopes & Oudeyer, 2016; Kidd & Hayden 2015; Marvin & Shohamy, 2016). A better understanding of curiosity might lead to further understanding of the mechanisms that drives people towards, and astray from, information and learning. According to present theory (Gottlieb et al., 2016; Kidd & Hayden 2015; Marvin & Shohamy, 2016) the opportunity for information gain is thought to spark curiosity and drive exploratory behaviour by being intrinsically motivating. Behaviour is extrinsically motivated when agents perform actions to achieve aims that are separate from their current endeavour, e.g. writing a thesis not because the inherent enjoyment of it (which would exemplify intrinsically motivated behaviour) but as a means to an end, e.g. as an instrument to get a job or motivated by dodging upcoming negative consequences. (See Oudeyer and Kaplan, 2007 for a conceptual discussion on intrinsic motivation). To this day few experimental studies have examined how, and if, information that is not externally motivated can drive exploratory behaviour, and importantly, no study to this day have seriously investigated if repeated exploratory behaviour driven by intrinsic information gain trades off against a comparable externally motivated task. If an intrinsically motivated task (a task that results in no other rewards but information gain) would be preferred to a comparable externally motivated task (which yields no information gain, but external rewards), then this should be indicative of the existence of a basic motivation towards information. The present study aims at putting the theory of intrinsically motivating information to a proper test, a test that requires an experimental task that captures the continuous and iterative nature of exploration, and to makes it possible to compare exploratory behaviour that is, and is not, externally motivated.

A review of the current literature on curiosity will now follow. It seems plausible that curiosity has evolved to help organisms discover regularities in a changing environment. Learning what there is in the world and to predict events is of benefit to species. A crux however is to understand how curiosity guides short term exploratory behaviour so that it can be advantageous in the long term (Kidd & Hayden, 2015; Gottlieb et al., 2016). Much of the present research on curiosity have been inspired by, and builds on, research done in the 1960’s and 70’s by Canadian psychologist Daniel Berlyne, this motivates a closer look at his theories. Berlyne thought that “we appropriately speak of ‘curiosity’ whenever a person or animal undergoes subjective uncertainty and is, in consequence, impelled to engage in specific exploratory responses” (1978, p.160). That exploratory behaviour occurs when the arousal level is supraoptimal, and that a moderate arousal potential is maximally rewarding (Berlyne, 1960). Berlyne coined the term perceptual curiosity, which refers to a type of curiosity that motivates behaviour that relieve uncertain perceptions. In humans, this curiosity can often coincide with epistemic curiosity – curiosity that have the potential to result in knowledge, which Berlyne defined as “information stored in structures of symbolic responses” (1978, p.144). However, the conceptual distinctions between different types of curiosity is debated (Kidd & Hayden, 2015), and to view curiosity as distinct from other psychological phenomena might not be very helpful. For instance, many aspects relevant to curiosity have been investigated in psychology departments under headings such as learning, exploration, attention, and play. Litman et al (2004), construed a Perceptual Curiosity questionnaire and identified perceptual curiosity as a personality construct related to both epistemic curiosity and sensation seeking, which suggest that perceptual curiosity involves a mix of seeking both knowledge and
sensory experience (this research does however only consider humans self-rated experience of curiosity and not actual observations of behaviour). In hope of side-stepping Wittgenstein’s view that “in psychology there are experimental methods and conceptual confusion” (p.232) we now turn to the most promising alternative in contemporary research, which is that information gain is intrinsically motivating.

A simpler explanation that transcends the earlier, and possibly unnecessary, distinctions is that curiosity is driven by information gain, and that information can be a driving force since it is intrinsically rewarding. This conceptualization is emphasized in a review by Gottlieb, Lopes and, Oudeyer (2016) and in an article by Marvin & Shohamy (2016). What it means for information to be intrinsically motivating is that the information gain does not need to be followed by a subsequent primary reward such as food, money, or water to result in reinforced learning. Support for this notion comes from imaging studies which have shown links between anticipation of information rated as highly curious and caudate structures known to be related to reward anticipation (Kang et al, 2009; Jepma, Verdonschot, Van Steenbergen, Rombouts & Nieuwenhuis, 2012). Further support comes from results showing that humans are willing to pay a higher price for information that they have rated as highly curious (Kang et al, 2009). Also, our non-human primate relatives often chose information about outcomes of gambles in favour of primary rewards (Blanchard, Hayden, & Bromberg-Martin, 2015), findings consistent with human subjects (Brydevall, Bennet, Murawski & Bode, 2017). Another idea is that the value of information depends on not on the amount of information, but rather on the expected information prediction error, which means that it is the gap between expected and gained information that is rewarding and results in learning (Marvin & Shohamy, 2016). Loewenstein (1994) put forward a theory that stated that curiosity was driven by agents’ predictions regarding the extent they thought the information-seeking would resolve uncertainty, which equals to satisfying curiosity. Information gain is not thought to reside only in humans’ perceptual or epistemic domains, it is thought to guide the exploratory behaviour of many different animal species. The difficulty however is to measure the information gain that seem to drive curiosity. Humans handle information with their senses, one of them being sight, which makes a good candidate for the study of information gain. Earlier research has studied the relation between curiosity and visual information but not properly quantified information, nor investigated individual differences in behaviour.

Experimental findings on “perceptual curiosity” have shown that humans generally prefer to watch a clear version of an image after having been presented with a corresponding blurred version of the clear image, compared to be presented with a clear but unrelated picture after presentation of a blurred image (Nicki, 1970). Presentation of blurred images also evoke longer EEG desynchronization compared to clear versions of blurred images, an effect depending on whether the blurred image was unknown to the observer (Berlyne, 1968). A more recent study by Jepma et al (2012) showed that perceptual uncertainty induced by blurred images, activated anterior insula and the anterior cingulate cortex, structures related to arousal and conflict, respectively. The relief of perceptual uncertainty (and curiosity) via reduction of blurred images activated striatum, an area related to reward processing. Taken together, these findings support the notion that curiosity is related to increased arousal and uncertainty, and that the regulation of arousal by removing uncertainty about the specific image that induced curiosity, is rewarding, and preferred over mere avoidance of the situation without new knowledge. Although, one should be cautious of informal reverse inference from activation of a brain region to a mental process (Poldrack, 2011), it still supports the notion that curiosity is related to mechanisms of arousal, uncertainty and reward.

In the field of visual search images, images are thought to have an inherent information distribution which affects visual attentional behaviour. Visual saliency refers to elements made
conspicuous by local spatial discontinuities in their properties of orientation-, intensity- and color information. These properties exert considerable bottom-up attentional force on the primate visual system (Itti & Koch, 2001). Examples of objects with salient qualities might be a green object surrounded by similar black objects, an object rotated 45 degrees compared to other objects, or an object with higher luminance compared to others. Distraction by saliency can delay response time and effectiveness of finding target during visual search (Foulsham & Underwood, 2009). According to a review by Wolfe and Horowitz (2017) five factors are known to together guide attention during visual search: bottom-up and top-down, scene-guidance, perceived value of features and the history of prior search. Also, color, motion, orientation and size are undoubt able guiding attributes. Results from Henderson et al (2009) indicate that during cognitively driven, top down, search in real world scenes (photographs) salient regions was not often visited. When searching for a target letter, a single colored pop-out distractor is possible to prevent distraction only if both the target and distractors are known.

If the salient features are predictable top-down factors can supress distraction (Theuwes Burger, 1998). Another bottom-up property affecting curiosity could be what Berlyne refers to as collative variables – properties of novelty, change, complexity, surprisingness and uncertainty (1960). Further, the image category can affect visual exploration of images (Kaspar & König, 2011). To summarize findings from the visual search and visual attention literature it is apparent that the information attended to depends on multiple factors.

There seems to be an abundance of variation in what visual information human beings crave. What is interesting for one individual is not necessarily interesting to another. People differ in their own self-ratings of perceptual curiosity (Litman, 2004). Such ratings can correspond to observable differences in individual’s behaviours towards visual information. For instance, self-rated measures of perceptual curiosity have been positively related to increased number of fixated regions in an image (Risko, Anderson, Lanthier, & Kingstone, 2012). However, research indicate that inter-individual differences in viewing behaviour are lower when images are subjectively rated as more interesting, suggesting that visual exploration is to a greater extent guided by image factors, when they are perceived as interesting (Kaspar & König, 2011). This suggests that a method that tries to understand curiosity must allow for individual differences.

Another relevant perspective on curiosity is its thought opposite: Boredom, which arises in situations of low information intake. Boredom can be understood as an adaptive signal that tells an organism to switch behaviour and manage the trade-off between exploration and exploitation, i.e. to look for more valuable opportunities or perform behaviour that result in well-known rewards. A study by Geana, Wilson, Daw, & Cohen (2016) showed positive relations between self-rated boredom and contexts with little useful information available, and that higher ratings of boredom were related to more exploratory behaviour. As most of us are familiar with, if the chain of events is highly predictable such as when performing monotonous tasks, the aversive state of boredom can appear. Putting these findings together it seems like boredom might be what precedes behaviours we understand as driven by curiosity in some situations, and signal when it is worthwhile to try to gain more information.

To get information, agents must in many situations engage in exploratory behaviours that result in information gain. In the common experimental paradigms researching curiosity, the presented information often consists of trivia questions presented by a procedure where the questions are shown to participants while they passively wait for answers to be presented to them (Baranes et al, 2015; Gruber, Gelman, & Ranganath, 2014; Kang et al, 2009; Shohamy & Marvin 2016), the same goes for studies regarding perceptual curiosity by Jepma et al. (2012) and Berlyne (1966). Earlier mentioned procedures only consider the effects of being exposed to perceptual uncertainty and making a choice whether to resolve uncertainty by pushing a
button, which of course is behaviour, but might be closer to decision-making regarding one’s curiosity, or identifying the neural correlates to the subjective experience of anticipation, rather than curiosity relieved by active exploratory behaviour. It could be something quite different when humans must continuously explore to gain information.

Vision is immensely important for humans and many other animals who move their bodies in the physical environment to assimilate and accommodate surprising, or not yet categorized, stimuli in their cognitive structures to know how to best respond. Investigating visual information gain, which in general from is a capacity we share with many other animals, and which does not require the use of language, could result in better comparisons of curiosity across humans of different ages, and with other animals.

If exploratory behaviour driven by curiosity is intrinsically motivating, the opportunity for visual information gain should be distracting to humans. Hence, when the same type of exploratory action is required to relieve curiosity and solve a given task, the conflict between them is expected to produce a mix of explorations aimed at relieving curiosity and complete a given task.

The aim of this thesis was to investigate, at the individual and group level, if the opportunity for visual information gain is distracting during an externally motivated visual search task, and to investigate if explorations that yields visual information would increase as a function of time. It was hypothesised that: (1) visual information with no instrumental value would induce distraction from an extrinsically incentivised visual search task; (1a) that there would be differences between background images regarding how distracting they were and that (1b) images independently rated as more interesting would be more distracting. (2) It was also hypothesised that visual information gain would increase as a function of time over (2a) the course of the experiment and (2b) within trials.

To address the stated hypotheses, an “image reveal”-paradigm was used. Participants were initially presented with a blurred image, by left-clicking the image a box of a more resolved portion of the image was revealed, and yet another click on same location revealed the full resolution of the image at that location. To get the visual information necessary to understand a single image, repeated exploratory behaviour was required by clicking the image. Active exploration would be an improvement over earlier studies (as suggested by Jepma et al., 2012). And since each exploratory click leaved a footprint as an image coordinate pair, precise measurement of behaviour was possible. The extrinsically motivated visual search task was designed such that a specific pattern of exploration was necessary to find the target. The specific pattern was constructed to prevent participants from getting the visual information necessary to understand what the images depicted, thereby constructing a situation of conflict between explorations needed to find the target and explorations needed for visual information gain. Finding the target progresses the test faster by decreasing the overall session time, which further incentivises the search task. To understand if it was visual information that was distracting, and not random haphazard clicking, two tasks where developed to produce individual topographical information-maps that located areas containing visual information needed to understand the image, similar to “interestingness maps” used by Onat, Açıkgöz, Schumann and König (2014). It is well known that top-down modulation of visual task influence eye-movements, e.g. performing visual search in an image and looking at an image with the intent to remember it yields different eye movement-behaviour (Castelano, Mack & Henderson, 2009). Since it is unknown how individuals extracting visual information when curious, two tasks were assumed to represent curiosity-driven exploration: one instructed participants to explore the image to understand it’s main theme. The other task was to explore the image to understand how interesting other observers would find the image. In the first task semantically meaningful objects have to be identified, in the second it is not explicitly required,
both are still thought to make participants explore regions they would have explored if being curious. Humans can find interesting locations in images even if the images do not depict anything semantically meaningful, such as fractal images (Onat, et al., 2014). The visual search task could then be compared to the exploratory pattern in main theme task and general interest task, which made it possible to quantify an information value for each click in the visual search task with respects to the targets location, the image main theme and the image general level of interest. Since transfer-effects were anticipated between the main theme task and general interest task to the visual search task; repeated exposure to stimuli can affect participants visual exploration of images, by decreasing saccade length and frequency, and increasing fixation durations (Kaspar & König, 2011).

**Method**

**Participants**

10 healthy participants (six females, four males, \( M = 26.10, SD = 4.93 \)) with normal, or corrected-to-normal, vision volunteered to participate. Participants were recruited by advertising on billboards around the Umeå University campus, by word of mouth or by advertising on student social media groups. All participant gave written consent before participation and received 150 SEK in return for their participation. Data collection was approved by the ethics board at Umeå university and in concordance with ethical guidelines from the Swedish Research Council (2002).

**Instruments and Materials**

Tasks were performed on a desktop computer, collection of responses and deliverance of stimuli was executed by custom made scripts in Matlab R2010 32b and Psychtoolbox (Brainard, 1997). Eye movements were recorded by an EyeLink 1000 (SR Reasearch Ltd.) and a head support was used to ensure more reliable eye-tracking. Measures of oculomotor behaviour were not used in the present study.

**Stimulus.** The stimulus used in this study were photographs that depict scenes of nature, animals, humans, man-made objects, events etc. All images level of interest was independently rated on a scale from 1-7 by a sample (\( N = 70 \)) via an Amazon Mechanical Turk experiment (MTurk), higher number indicated a more interesting image. For this study, the 40 lowest and 40 highest rated images from a set of 262 images were chosen, see Figure 1 for example images. Stimuli were presented at 75 cm distance on a 532 x 299 mm LCD monitor with 1920 X 1080 resolution and a pixel size of .27 mm. Stimulus image sizes varied between 686 X 1080 pixels and 1344 X 756 pixels. Each image was generated and processed anew before each trial.

The selection and processing of stimulus to this study has been developed as part of a curiosity-project lead by Linus Holm (Umeå University) and Paul Schrater (University of Minnesota). The selection of images, development of the “image-reveal” task and image-processing was developed in Schrater’s lab at the University of Minnesota. The assembly of the experiment was coded and implemented by Linus Holm and Gustaf Wadenholt in Linus Holm’s lab at Umeå University. The author was responsible for collecting, analysing, and visualising the data material in this thesis.
Procedure

Participants were seated on a height-adjustable chair at a table with adjustable height to ensure a comfortable seating position. Participants performed a total of 240 trials arranged in 3 blocks. The first block consisted of 80 trials of a visual search task which was followed by two blocks of exploratory tasks. A block of trials was completed in approximately 20-25 minutes, all blocks of trials were performed under one session with a five-minute break between trials, resulting in 70 to 85 minutes to complete all exploratory tasks. After the visual search task five minutes were reserved to complete a questionnaire. Participants were also briefly interviewed and debriefed after they had completed tasks and questionnaires. Behavioural responses collected were the xy-location of the cursor at the time of a mouse click.

Participants were first given the visual search task which consisted of finding the target object, consisting of the cartoon figure Waldo, who was hidden under 2 layers of blurred images, and blended to a third clear image. The blurred image did not initially give any information of the target. Waldo was present in 20% of the images at a random location on a circle with a fixed radius. Before the block of search trials started, participants were presented with a demo that explicitly showed all possible locations where Waldo could be found. Visual search is an incentivized task since participants were instructed by the experimenter to look for Waldo, and additionally, they received feedback if Waldo was found. Finally, if the participant found Waldo, this progresses the test faster by decreasing the overall session time. Participants got no “external” or instrumental rewards from satisfying curiosity about the picture. Prior to the search task participants performed at least eight training trials where Waldo was present, ensuring that the participants had understood the possible locations where Waldo could be placed (see Figure 2). A trial was completed either when the target was found, after 10 seconds had elapsed or after 50 exploratory clicks had been executed. Participants indicated that they found the target by first pressing space and then left-clicking on the target. If waldo was not found in 10 seconds, participants were again presented with the image and could take a guess by clicking or press space to indicate that Waldo was not in the picture. Participants received feedback if they found Waldo. After the visual search task was completed, participants

Figure 1. Background images. Row (A) initial blur level. (B) Example of exploration during main theme task or general interest task. (C) Clear images. (D) Example of exploration during visual search.
completed the main theme task and general interest task in counter balanced order across participants.

In the main theme task participants were instructed to explore the image to identify its main theme and then to pick the best answer option out of a seven alternative forced choices (7AFC). Participants had 10 seconds to explore the image and 5 seconds to pick an answer. The 7AFC alternatives were developed by showing the images, in different levels of pixilation, to five students. The students were asked to generate hypotheses regarding the main theme in the pixelated images which was used to identify reasonable response alternatives for the 7AFC task. If some pictures did not get enough alternatives, the author of this thesis estimated possible alternatives. The correct answer of 7AFC was subsequently validated by the students. Most of the correct answers to the 7AFC were the category of the main object depicted in the image (e.g., animal, flower, etc). If the image depicted an event, the right answer could be “feeding a bird” etc.

Figure 2. Flowchart over experimental design. (A) Visual search. A trial consists of a fixation cross for .3 s followed by 10 s of search or 50 clicks. If the target is found and indicated within 10 s or 50 clicks, the participants receive feedback and proceed to the next trial. Else, if the target is not found, the explored image is presented again, and participants can either guess the targets location or indicate target as missing. Participants received feedback and proceeded to the next trial. After 80 trials participants proceeded to either (B) main theme task or (C) general interest task.
In the general interest task, participants were instructed to explore the image to understand how interesting it was for other people, and to estimate the mean interestingness-rating of the picture. The correct answers to the interestingness of pictures were based on a composite median score from the MTurk experiment. The images’ level of interest were rated in different levels of blur in the MTurk experiment.

The reveal box was in all three tasks set as a square with side length of 60 pixels centred on the mouse cursor.

After participants completed the search tasks they filled out a personality measure. The personality data was not used in the present study.

Data Analysis

In the visual search task each mouse klick was stored as an xy-coordinate that mapped onto the image. Each click during trials was given an information score by calculating how well the click overlapped with the map of possible Waldo locations and maps over regions needed to complete the main theme task and general interest task (see Figure 3). These three different types of information are henceforth denoted *Waldo Information* (W.info), *Theme information* (T.info) and *Interest information* (I.info), respectively. Z-scored T.info and I.info was calculated for each participant on each trial by comparing the participant’s mean T.info and I.info during one trial, to a sampled mean score. The sampling was done by randomly drawing a coordinate from the participant’s pool of empirical clicks during all visual searches’ (about 3000 clicks per participant). The sample coordinate was then given a T.info or I.info by comparing it to the participant’s unique model of how they explored the image during the main theme task and interest conditions, respectively. 50 n = 1000 samples were randomly drawn, and the mean scores were computed for each participant’s trial. The mean empirical score for each trial was then given a p-value by calculating the proportion of sampled means falling under the empirical mean. The p-value was then converted to a z-score. This process makes the random sample reflect individual differences in search performance and make participant functions as their own controls, which takes heed of each participant’s individual way of exploring an image to gain visual information.

![Figure 3](image_url)

*Figure 3.* First row, visualizations of empirical exploration from one participant in one trial. (A) empirical exploratory clicks in visual search, (B) main theme and (C) general interest. (D) trial stimulus. Second row, visualization of the maps used to calculate information scores. (E) Map of where Waldo can be found (same for all participants), followed by unique models based on participants behaviour in the main theme task (F), and general interest (G). The empirical xy-clicks from visual search were evaluated against the three maps in the second row resulting in the participants W.info, T.info and I.info quotient. Which basically amounts to putting (A) over (E) and (F) and (G) and calculate a quotient.
Results

Descriptive statistics

Participants made a total of 101830 exploratory clicks during task 1 ($M = 3301.8$, $SD = 224.42$), task 2 ($M = 3440.6$, $SD = 153.12$) and task 3 ($M = 3431.6$ $SD = 278.14$). The mean number of exploratory clicks within trials were 42.43. See table 1 for performance scores in task 1 and 3.

The mean raw scored $W$.info was $3.46 \times 10^{-6}$ ($SD = 6.49 \times 10^{-7}$) mean raw scored $T$.info $= 1.81 \times 10^{-6}$ ($SD = 1.34 \times 10^{-7}$) and mean raw scored $I$.info $= 1.86 \times 10^{-6}$ ($SD = 1.34 \times 10^{-7}$). See Figure 4 for information levels from trial 1 to 80. See Figure 5 for raw information levels from clicks 1:50 within trials.

The independently rated interest-scores of images ranged from 1 to 5 ($M = 2.95$, $SD = 1.11$). The group of 40 highest rated images had a mean of 3.94 ($SD = .46$), and 40 lowest a mean of 1.96 ($SD = 0.55$).

Figure 4. Plot of aggregated mean scores over trial 1 to 80. Solid lines represent mean raw information value and shaded areas 95% confidence intervals.
Table 1

<table>
<thead>
<tr>
<th></th>
<th>Visual Search</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Waldo present</td>
<td>Waldo absent</td>
<td>Sensitivity</td>
<td>GIT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hit</td>
<td>Miss</td>
<td>CR</td>
<td>FA</td>
<td>d’</td>
</tr>
<tr>
<td>1</td>
<td>.63</td>
<td>.38</td>
<td>0</td>
<td>0</td>
<td>NA</td>
</tr>
<tr>
<td>2</td>
<td>.81</td>
<td>.19</td>
<td>.92</td>
<td>.05</td>
<td>2.52</td>
</tr>
<tr>
<td>3</td>
<td>.50</td>
<td>.50</td>
<td>.39</td>
<td>.03</td>
<td>1.88</td>
</tr>
<tr>
<td>4</td>
<td>.63</td>
<td>.38</td>
<td>.74</td>
<td>.03</td>
<td>2.21</td>
</tr>
<tr>
<td>5</td>
<td>.94</td>
<td>.06</td>
<td>.98</td>
<td>.02</td>
<td>3.61</td>
</tr>
<tr>
<td>6</td>
<td>.63</td>
<td>.38</td>
<td>.13</td>
<td>.03</td>
<td>2.21</td>
</tr>
<tr>
<td>7</td>
<td>.81</td>
<td>.19</td>
<td>.58</td>
<td>.03</td>
<td>2.76</td>
</tr>
<tr>
<td>8</td>
<td>.75</td>
<td>.25</td>
<td>.31</td>
<td>.03</td>
<td>2.56</td>
</tr>
<tr>
<td>9</td>
<td>.50</td>
<td>.50</td>
<td>.92</td>
<td>.13</td>
<td>1.13</td>
</tr>
<tr>
<td>10</td>
<td>.69</td>
<td>.31</td>
<td>.95</td>
<td>.08</td>
<td>1.90</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>.69 (.14)</td>
<td>.31 (.14)</td>
<td>.59 (.37)</td>
<td>.04 (.04)</td>
<td>2.31 (.68)</td>
</tr>
</tbody>
</table>

*Note. main theme task results are missing in table 1 due to technical difficulties. For what it’s worth, my own observations of participants behaviour during testing, as well as post-test interviews, gave me the impression that participants had been focused on the task and did not find it difficult. I expect that participants had high proportions right answers on the main theme task.

*p < .05.

Figure 5. Plot of aggregated mean scores clicks within trials. Solid lines represent mean raw information value and shaded areas 95% confidence intervals.
Inferential statistics

**Individual level analysis.** To investigate if individual participants were distracted by informative regions of the images and which images that were distracting, participants z-scored T.info and I.info was compared to the cut off for statistical significance which was set to \( p = .05 \), thus z-scores above 1.96 are understood as observations not likely observed by random clicking. See Figure 5 for a visualization of the results.

To investigate if images with a higher interest level, resulted in higher information gain than lower rated images, paired samples \( t \)-tests were conducted between z-scores on stimuli grouped in high- and low interest conditions, consisting of the 40 highest rated images versus the 40 lowest rated images. See Table 2.

![Figure 6](image.png)

*Figure 6.* Illustration over participants level of distracting information gains. Upper figure refers to distraction towards T.info, lower figure image distraction towards I.info. Purple and green values have z-scores above 1.96 which is equivalent to \( p < .05 \). \( M \) = mean information gain from participants 1:10. The blue values in each figure give information on the variability of the z-values below cut-off with light blue and dark blue values indicating lower T.info- and I.info respectively. White coloured boxes represent missing data.
Group level analysis. To investigate if images were generally distracting, participants mean z-score was evaluated. Mean z-scored T.info = .23 (SD = .11) and mean z-scored I.info = .16 (SD = .52). Two stimuli resulted in z-scored T.info meeting the p = .05 threshold, namely stimuli 28 (z = 2.07) and stimuli 62 (z = 2.56). Image 62 also had a z-scored I.info above cut-off (z = 2.79). Mean z-scores are represented at the end rows in Figure 5. See Figure 7 for topographical information maps based on aggregated exploration in tasks in conditions of image 28 and 62.

<table>
<thead>
<tr>
<th>Participants</th>
<th>z-scored T.info</th>
<th>z-scored I.info</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.12</td>
<td>-1.54</td>
</tr>
<tr>
<td>2</td>
<td>-0.66</td>
<td>-0.71</td>
</tr>
<tr>
<td>3</td>
<td>-0.64</td>
<td>-0.80</td>
</tr>
<tr>
<td>4</td>
<td>-0.25</td>
<td>0.40</td>
</tr>
<tr>
<td>5</td>
<td>0.54</td>
<td>0.18</td>
</tr>
<tr>
<td>6</td>
<td>0.46</td>
<td>0.64</td>
</tr>
<tr>
<td>7</td>
<td>-0.44</td>
<td>1.00</td>
</tr>
<tr>
<td>8</td>
<td>-0.41</td>
<td>-0.92</td>
</tr>
<tr>
<td>9</td>
<td>-0.51</td>
<td>-1.23</td>
</tr>
<tr>
<td>10</td>
<td>2.13</td>
<td>0.61</td>
</tr>
<tr>
<td>M</td>
<td>-0.06</td>
<td>-0.92</td>
</tr>
</tbody>
</table>

Note. df are under 39 for some participants since they did not complete all trials. Missing values are excluded from analysis.

*p<.05.

**Table 2**

Results from paired samples t-tests between high- and low interest conditions

<table>
<thead>
<tr>
<th>Participants</th>
<th>z-scored T.info</th>
<th>z-scored I.info</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.12</td>
<td>-1.54</td>
</tr>
<tr>
<td>2</td>
<td>-0.66</td>
<td>-0.71</td>
</tr>
<tr>
<td>3</td>
<td>-0.64</td>
<td>-0.80</td>
</tr>
<tr>
<td>4</td>
<td>-0.25</td>
<td>0.40</td>
</tr>
<tr>
<td>5</td>
<td>0.54</td>
<td>0.18</td>
</tr>
<tr>
<td>6</td>
<td>0.46</td>
<td>0.64</td>
</tr>
<tr>
<td>7</td>
<td>-0.44</td>
<td>1.00</td>
</tr>
<tr>
<td>8</td>
<td>-0.41</td>
<td>-0.92</td>
</tr>
<tr>
<td>9</td>
<td>-0.51</td>
<td>-1.23</td>
</tr>
<tr>
<td>10</td>
<td>2.13</td>
<td>0.61</td>
</tr>
<tr>
<td>M</td>
<td>-0.06</td>
<td>-0.92</td>
</tr>
</tbody>
</table>

Pairing sample t-tests were conducted to compare information scores from high- and low interest conditions. No significant difference was observed between W.info in high- and low conditions t(39) = 1.76, no significant difference was observed between T.info in high-
and low conditions $t(39) = 1.87, p > .05$, and no significant difference was observed between I.info in high- and low conditions $t(39) = 1.84, p > .05$.

Correlations between information scores and the median interest level of images 1:80 were calculated to investigate the relationship between participants information gain and the previously measured interest-ratings of background stimuli. No significant correlation was observed between the images interest-rating and mean W.info $r(78) = .26, p > .05$, T.info, $r(78) = .14, p > .05$; I.info, $r(78) = .16, p > .05$). See Figure 8.

\[ \text{Figure 8. Correlations between information gain and independent interest ratings of images. Plots from left to right: W.info, T.info and I.info.} \]

To investigate if visual information gain increased over trials, trial order was correlated with visual information gain, see Figure 9. There was a non-significant positive correlation between trial order and W.info, $r(78) = .18, p = .11$, and non-significant positive correlation between T.info and trial order, $r(78) = .08, p = .45$. There was however a significant positive correlation between trial order and I.info, $r(78) = .25, p = .03$.

\[ \text{Figure 9. From left to right: correlation between trials ordered on time of presentation and W.info, T.info and I.info, respectively. *p < .05.} \]

To investigate if visual information gain increased over time within trials, correlation was calculated to assess the relationship between click order and information gain. There was a very strong significant negative correlation between click order and W.info $r(48) = -.86, p < .05$. A non-significant positive relationship between click order and T.info, $r(48) = .21, p > .05$, as well as click order and I.info, $r(48) = .27, p > .05$). See Figure 10.
The aim of the present study was to investigate if the opportunity of visual information gain could drive participants away from an extrinsically motivated task, and if the hunger for visual information would increase over time. Briefly, the results suggest that there are instances where visual information gain can distract participants away from an extrinsically motivated task, there is however much heterogeneity between participants regarding which stimuli appear distracting. The background images “objective” interest level, as assessed from the MTurk experiment, did not affect curiosity as expressed in clicks on image-informative regions. Results also suggest that extraction of visual information increased over trials, but not within trials. These results are further discussed below.

From Table 1, we can observe that most participants had high $d'$, and therefore with confidence assume that participants did try to find the target during the visual search task, and they could reliably discriminate when the target was present and not present. Unfortunately, performance data is missing for the main theme task, but I will assume that participants were focused on that task as well. And even though proportions of right answers on the general interest task were generally low, most of the participants performed better than what would be expected by chance, and many participants experienced difficulties in estimating the “objective” interest levels of images, I will assume that participants tried their best at this task as well. Considering these empirically supported assumptions of participants engagement, we have reason to proceed in examining how the results relate to the main hypotheses.

Generally, on almost all trials, participants had higher W.info compared to T.info and I.info, which means that participants were predominantly exploring the image with the intention of finding Waldo and adhering to the extrinsically motivated visual search task. Thus, the rewards of visual information gain could not compete with those from the extrinsically incentivised task. In other words, participants were not generally so curious about the image that they abandoned the instructed task at hand. These findings are not incommensurable with present understanding of curiosity as driven by primary rewards in the form of information gains, as suggested current researchers of curiosity (Brydewall et al, 2017; Gottliebet al, 2016; Jepma et al, 2012, Marvin & Shohamy, 2016). However, the trade-off between intrinsically motivated and extrinsically motivated exploration is not comparable in this task. And as will be discussed in limitations, it is probably not the case that the visual search task is only extrinsically driven.

That the general finding, that images were not generally curiosity evoking, can be explained by variabilities regarding which images that individuals where curious about, as can be seen in Figure 6. This finding is not surprising considering our everyday knowledge that
people are interested in different things and are in line with results from Litman (2004) and Risko et al (2012). There were however two images that stood out as generally curiosity evoking: image 28, a flower, drew significantly higher T.info, and image 62, a grasshopper, drew both higher T.info and L.info than what would be expected by random sampling. An ocular examination of Figure 7 show that there is much overlap between behaviour in visual search and both the main theme task and the general interest task regarding these two pictures. How come participants were generally demonstrating exploratory behaviour towards visual information in these regions in these images? Individuals information-maps overlapped in most images, so it seems like participants differ in what they are interested in, but not how they are interested, the observation that there is little variation between humans which image-regions humans find interesting is consistent with results from Onat et al. (2014) who report overlap between which areas humans find interesting, and Kaspar and König (2011) who report less differences in inter-individual exploration of images subjectively rated as more interesting. It also seems to be overlap between salient areas and explored areas in images 28 and 68, results also in line with Onat el al. (2014) who report some overlap between subjectively indicated interesting areas and saliency-computed areas. The flower had an interest rating of 2, and the image of the grasshopper an interest rating of 4, which is similar to other images in the set. The collative variables of these two images were not especially high compared to other images used; they do not appear as more complex, novel, surprising or uncertain than other images. One thing that characterizes both images and make them different than most other images was the uniform background surrounding the semantically meaningful object. Exploration of these background regions does not result in any information gain at all. There are at least five other images with similar uniform backgrounds, but they were not significantly distracting. What differentiated the two distracting images from the other images with uniform backgrounds was the spatial position of the visual object depicted; in the two generally distracting images, there was not much visual information coinciding with Waldo’s possible position. That is, participants did not gain any information about the content at all by searching Waldo. In all other images Waldo’s possible placement coincide with informative regions, resulting in small amounts of “free” T.info and L.info gain during the visual search task, since the exploratory path needed to find Waldo coincides with more dense areas of visual information. In the images depicting the grasshopper and flower, the possible Waldo-locations lie in regions where no information is to be gained, which necessitates clear deviances from the search pattern to gain any visual information. I think one can, without a formal analysis, be quite certain that the high-level visual T.info and L.info maps coincide considerable with visual salient areas of the two images, and since visual saliency in many cases can explain considerable amounts of attention and at the same time coincides with interesting image regions, which also have been shown to draw attention (Onat et al, 2014) there is no conflict between the images low-level saliency properties and high-level interestingness regions. One could think that saliency effects and visual information work together to draw curious explorations towards the main objects. Seen from the perspective of Loewenstein’s (1994) theory that it is the extent to which predicted information gain resolves uncertainty, these two images are quite easy to reduce uncertainty of, especially the grasshopper, since you only need to perform a small number of clicks to reduce a lot of uncertainty regarding the visual information in the image, which makes this image stand out as more accessible and thus more curiosity evoking.

At group level, results from the paired samples t-test between high- and low interest conditions, showed that images independently rated as more interesting were not more distracting than others, and the amount of information gains from images was not significantly correlated to their independent interest rating values. This is surprising, since the independent ratings suggest that some images are “objectively”, or generally, more interesting than others.
Results indicate great variability between what people find interesting, see Figure 6. However, with a sample size of N = 10 the sample could consist of a cluster of individuals who are not representative enough to warrant general conclusions. Also, there is much inter-individual differences in the exploratory behaviour measured in this study, which set in relation to results from Kaspar and König (2011), which showed that lower interest levels were related to greater variability, would suggest that most images in the present study was not that interesting. In post-test interviews, several participants had objections about the “objective” interest ratings, not agreeing with them and one participant even felt sad because the ratings were generally very low.

At the individual level, only one participant was significantly more curious about images categorized as highly interesting, as measured by T.info gain compared to W.info gain. It might be the case that this participant held similar opinions about the interestingness level of images as the independent MTurk sample, this curious participant was also better at estimating the correct independent interest levels, compared to eight other participants. The observation that for this participant high-interest images drew more T.info, but not, I.info, compared to random sampled search clicks, is an interesting result, since the impression is that exploration with intent of finding the main theme and evaluating interestingness at the aggregated level seem to yield similar information extracting behaviour, as seen on topographical information maps. There may be significant differences between the way one explores an image based on task, as seen in eye-tracking studies (Castelano et al, 2009). However, this could also be explained by the participant being curious about the image, such that the participant, on a second presentation of the image remembers parts of the image and wants to uncover areas of the image that are still informative. There are reasons to expect transfer effects between tasks since participants are exposed to the same images. One participant mentioned memory effects such that, areas explored in the previous task, were avoided in favour of areas not explored in the previous task. This possible confound could be assessed and possibly corrected for by constructing a composite score of both T.info and I.info, these analyses are however due to time constraint limitations pertaining to the author, beyond the scope of the present study. It might also be the case that there exist such great differences between individual’s subjects of curiosity, that it is not meaningfully related to larger trends, there should however be clusters of groups with similar interests who exhibit similar curiosity-driven behaviours. Larger studies are needed to properly evaluate the inter-individual variability, and which categories that group people together. Perhaps, personalities may account for interest distributions in the task. An interesting study would be to relate personality measures with visual exploration of images with different image content categories.

In the more detailed analysis of time effects, there was a small, but significant, increase in I.info gain over trials, but no reliable significant relationship between time and T.info and W.info. An interpretation of these results is that the participants over trials became more curious about the images (as measured by I.info). However, no reliable results were observed regarding the relationship between I.info and W.info, however the trend was positive. The results suggest that participants manage to learn how to satisfy both the extrinsic imperative to find Waldo and an at the same time satisfying their curiosity. This I think strongly suggests that visual information gain is rewarding, since it must be a much more difficult task gain more visual information without compromising performance on visual search, which supports the understanding of information as rewarding (Blanchard et al., 2015; Gottlieb et al., 2016; Jepma et al., 2012; Marvin & Shohamy, 2016). The result, indicating that participants tried to optimise information gain over time, is in line with results from Geana et al. (2016).

Results from time effects within trials show that participants performance on gaining W.info steadily decreased within a trial, but that there was no significant relationship between
time and visual information gain. Results on steadily decreasing W.info could be interpreted as indicating that participants became distracted from the task. However, T.info and I.info does not show a significant increase in visual information gain within trials and does not explain the distraction. My suspicion is that the declining W.info score is a function of the participants still trying to find Waldo, however due to uncertain memory representation of the circles diameter and position, they probably explored the outer areas of Waldo’s possible positions. It would be able possible to test this formally, but I make this assumption based on observation during testing and interviews with participants. Focusing on the last 5 clicks, there is a dramatical increase in visual information gain, and at the same time a dramatical decrease in W.info, suggesting that this might be a turning point regarding the exploration-exploitation trade-off (Geana et al., 2016) where participants are switching to their own formulated top-down task, probably to gain new information in any way. It would be interesting to extend the time and number of clicks so one could asses if the observed trend really is a switch towards increased information extraction. Should this be the case, it suggests that the driving force of curiosity quickly takes over and becomes a prioritized top down goal. Since this requires some energy, the internal neural system must find new information gain intrinsically worth the effort. Participants could equally likely just have had given up, stopped clicking and stop wasting energy. After all, it does take some energy to perform about 4 click/s during 80 consequent trials lasting 10 seconds each. This is consistent with findings that indicate that humans value new information, this interpretation must however be taken with caution, since all participants did not perform 50 clicks in each trial, either because they found Waldo, or were out of time, before 50 clicks had been performed. Clicks were not registered after Waldo had been found, so results are not influenced by any clicks performed after Waldo had been found.

Again, the two different kinds of information scores, and models of curiosity-related exploration, seem to tap into different behaviours. Understanding what a picture depicts is a more straightforward task than determining how interesting it is to others. However, it seems like it is not necessarily the case, that one need to know what the picture depicts to know if it is interesting, which one might intuitively think. Results from Onat et al. (2014) show that their similar “interestingness maps” were good predictors of oculomotor behaviour during free-viewing of urban images, complementing saliency-models of visual attention. However, fixation-patterns on natural images were more dissimilar than to urban images with respects to their “interesting-maps”. Suggesting that image category can have an effect of how well behaviorally indicated measures of visual information compares to information extraction by visual attention.

I think one are justified in suspecting that curiosity, so to speak, lurked under the surface during the visual search task, but was constrained by top-down prioritization. However, over time, participants accommodated to the situation of low information gain by learning how to better satisfy both the extrinsic imperative to finish the task, and the intrinsic imperative to gain new knowledge.

A limitation in this study is that it is difficult to know to what extent the visual search task was extrinsically motivating. It was the case that participants had extrinsic motivations of finding Waldo for task compliance and to reduce total time in the experiment, but they probably also had additionally intrinsic motivations towards finding Waldo. Most of the participants reported the visual search task as challenging and fun. Participants also became very frustrated when they did not find Waldo, which made them try even harder. This means that the visual search task was not completely, or maybe not even primarily, extrinsically driven. An alternative interpretation is that the different intrinsic motivation goals may have competed in the task. The threshold for a trade-off towards exploring new opportunities of rewarding information gains compared to exploiting the current task is thus probably set very high, and
might require some extremely interesting images, or a more boring task. Even if searching for Waldo might be a rather intrinsically motivated task, understanding the image was not associated with any instrumental value at all, making it possible to draw conclusions from the results.

Another possible in data interpretation is participants personality. Every psychology experiment is in a sense a possible social psychology experiment: participants were probably to different extent careful to not break the rules that the experimenter had given them, therefore, personality constructs, such as conscientiousness, might be a modulating factor regarding how much participants inhibited their curiosity. A further limitation in this study is the sample size. Flowers were overall not interesting to the independent sample responsible for the “objective” interest ratings, but it might be the case that there were a couple of botanists or entomologists among the participants in this study. Possible effects of individual differences with regards to how well the set of images match participants special interests are not controlled for. Behavioural idiosyncrasies regarding special interest contents within a single image should however have been handled well since participants scores are evaluated against their own information maps. But more probable, rather than effects of special interests, is that image categories like Urban, Natural (Kaspar & König, 2011; Onat et al., 2014), and maybe Animals, have modulated exploratory behaviour.

Yet another factor which could affect the patterns of visual explorations in this study are the varying images sizes used. When free viewing images, larger images typically yield increased oculomotor exploration and decreased exploitation (Gamiero, Kaspar, S.König, Nordholt & P.König, 2017), this study did analyse oculomotor behaviour, visual exploration has been shown to cooccur with behavioural indications of interesting image regions (Onat et al. 2014), there is reason to expect that image size could have affect the visuo-motor task in this thesis.

A suggestion for further research is to make a slight alteration of this task to make it more boring. It is possible to pitch an extrinsically motivated task more stringently by really preventing any coinciding visual information gain while searching the target. An idea to achieve this is to cut out a circle from the background image where Waldo can be put, performing the search would then not give any “bonus” information about the scene. At the same time, this design would solve the possibility of participants forgetting the Waldos possible positions, the task thus it becomes much less challenging and probably much more boring, and more extrinsically motivated. One should also let participants perform the task with more generous time limits and let participants control the time they spend on the task. They would then have more time available to explore the image if they wanted to, but to continue to the next trials they would have to find the target. Then a measure of the information value could be the time spent exploring the image, similar to how waiting time to receive answers have been used to measure level of curiosity (see Kang et al., 2009).

Evaluating the contribution of saliency and bottom up image properties would have been informative on how much curiosity was driven by these properties drawing visual attention. It would also be possible to augment the experiment with physiological measures of arousal to investigate the hypothesis of curiosity being arousal-regulating (Berlyne, 1960). Further studies should also evaluate the relationship between eye-movement exploration and behavioral exploration, measures of visual attention would be a very good complement.

Results from this study does support the theory that visual information is intrinsically rewarding by showing that participants align their behavior to gain as much information as they can over time, while at the same time maintaining their level of performance on a conflicting task. Furthermore, the results demonstrate that curiosity about visual information at times can exert enough force to distract participants from performing an extrinsically rewarding task. It
is however not clear what image-properties evoked participants curiosity. Results show that there is much variability in which images that evokes curiosity, and objective measures of interest level was not related to increased exploration as hypothesized. The presents investigation gives reasons to understand curiosity as the hunger for information gains, and that information can be seen as a primary reward in humans.
Reference list


