Written home exam in Cognitive Psychology in HMI

This document consists of general instructions, instructions regarding the use of CHATGPT, and home exam questions. Read the whole document.

General instructions

Download and save this document, answer the exam questions in this document. When you are done answering the questions, save the document as a pdf-file, and post the document in iLearn "Post your written home exam here".

In this home exam it is mandatory to use and refer to the course literature:

- ☐ Cognitive psychology and it´s implications, Anderson, and
- ☐ Understanding computers and cognition, Winograd & Flores
- ☐ Foundations of behavioural research, Kerlinger, or
- ☐ Social research methods, Bryman

In this home exam, it is allowed to use the following resources:

- ☐ ChatGPT
- ☐ Literature resources other than the course literature

Note that the use of CHATGPT in no way guarantees a correct answer, and could hence result in a lower grade, or a fail.

In this document, in case you choose to use CHATGPT, check corresponding box (below). If you have not used CHATGPT, check corresponding box (below).

If you choose to use CHATGPT, you need to adhere to and follow the instructions as presented below.

CHATGPT and AI tools in written home exam

During this home exam it is permitted to use CHATGPT and similar search/AI tools provided that you follow all of the instructions below. Ideally it will lead to more time for analysis, elaboration and discussion.

You will find the service here: https://openai.com/blog/chatgpt/

What is written below in relation to CHATGPT is applicable to all AI tools with similar functionality. If you choose to use CHATGPT, or a similar AI tool, you have to follow the instructions below.

1) You have to provide a printout of the question(s)/promts you used and the full text provided by CHATGPT that you used. This should be copied into this document.
   - ☐ It is very important which question/prompt you use in CHAT GPT, small differences in formulation(s) may generate very different quality and plausibility in the answers provided by CHATGPT. In effect, this means that CHATGPT does not provide you with the full answer to the home exam questions. Thus, you need to analyse the answer(s)

CHATGPT provides you with, and adjust the question(s)/prompt(s) where necessary, as well as the text/answers provided by CHATGPT. In other words, engage in an iterative process with CHATGPT.

- In your answer to the home exam questions, you have to highlight in yellow in your answer the parts that are verbatim or modified CHATGPT text. Further, in the answer(s) to the exam questions, you have to refer to the course literature with page reference (i.e., Anderson and/or Winograd & Flores, Kerlinger or Bryman) where applicable. You need to add your own text which should consist of elaborations of the CHATGPT text. This means you have to analyse the answers/texts that CHATGPT provides you with, make necessary adjustment and complement the text with references to the course literature.

2) You have to provide a description on how you have used CHATGPT.
   You should provide this in the first paragraph of your answers to the home exam questions. The explanation should be ½ page. This explanation should contain the following components:
   - Which questions/prompts resulted in CHATGPT-text/answers that you chose to use and elaborate? You need to shortly describe and motivate your choice and course of action.
   - If you find that the answers from CHATGPT are of poor quality and you have not used these at all, please state that and explain how you came to the conclusion that the answers were of such bad quality. This could for example be an explanation that based on your use of CHATGPT you found that the answers CHATGPT provided were wrong and/or inaccurate.

3) Plagiarism and Impact on Grading:
   a) The use of text from CHATGPT without indicating this (in dedicated checkbox below), mentioning it in your text or the explanation (section 2 above) and having used this tool is considered to be plagiarism/cheating and will immediately result into a formal complaint with Stockholm University.
   b) The use of text from CHATGPT where you mention in your text or the explanation (section 2 above) that you have used this tool but without adding references to the course literature where applicable, is considered to be bad quality which may lead to a lower grade, including a fail.
   c) CHATGPT may help you to write your answers to the home exam questions. However, you have to conduct your own analysis and text. The grading will consider to what extent you have done that.

Exam questions

This written home exam consists of two related questions, both of which are based on the group work you participated in during the course. For both of these questions, you must use and refer to the data, results and interpretations presented in the group report you participated in writing.

Question 1: In this part of the written home exam, as compared to what was presented in the group report, you must do an in-depth critical review of the theoretical grounding made in the work;

- Given the results the experiment generated, can you say with certainty that what was intended to study, really was what was studied? Grounded in collected data and interpretations of results, discuss and present arguments supported by the course literature that support that what was intended to be studied was indeed studied. Likewise, present arguments rooted in the course literature that may point to the opposite.

As related to question 1, question 2 is completed below as a separate answer. Therefore, answer question 1 in a separate section, then question 2. In question 2, references can be made to question 1.

Question 2: In this part of the written home exam, as compared to what was presented in the group report, anchored in data/results and course literature (method and theory), you must make an in-depth critical review of the implementation of the experiment. This includes;

- Definition of variables, i.e., dependent variable, independent variable(s) and control variables (note that there is a natural connection here to theory and question 1).
- The material produced/used as independent variable(s), (note that here too there is a natural connection to theory and question 1).
- As related to the two points above, results and interpretation of results.

General

Applies to both question 1 and question 2;

- In the critical review, using the argument that too few research subjects were included in the experiment does not constitute an argument as it was already included in the conditions for the experiment.
- In your answers, and presented arguments in question 1 and question 2, refer to course literature with page references.

Check the boxes below indicating whether you used/did not use CHATGPT in your answers to question 1 and question 2:

☐ I have used CHATGPT in my answer to question 1 of the written home exam, and also followed the instructions for how to use CHATGPT.

☒ I have not used CHATGPT in my answer to question 1 of the written home exam.

☐ I have used CHATGPT in my answer to question 2 of the written home exam, and also followed the instructions for how to use CHATGPT.

☒ I have not used CHATGPT in my answer to question 2 of the written home exam.

My answer to question 1

The conducted study, aimed to investigate how distractions (in this case advertisements on a webpage) might affect reader's attention to that text, and thus, making a statistical correlation between the two aforementioned variables. Although the results are not incontrovertible, they showed that the ads can influence the reader, but cannot distract completely and that they can, though, influence reader's short-term information retention ability. As for this review, an attempt to commend and criticize, by providing both supporting and opposing arguments rooted in the literature, will be made. And to give a brief answer before digging into a more detailed review of the literature, it is not clear that what was intended to study was indeed what was studied.

Starting with the attention, as already mentioned in the research, to measure a subject's attention, his eye movement was tracked, and specifically, by analyzing its pupil fixation using an eye-tracking software. This approach was adopted, because it is stated that '*a shift of attention often precedes the corresponding eye-movement'* (Anderson, 2009, p.70), and furthermore Wojdynski and Bang mention that '*eye movements, in the form of pupil fixations within a particular area, are indicative of allocation of attentional resources to that specific area*' (Distraction effects of contextual advertising on online news processing: an eye-tracking study, 2016, p.4). The results, then, show that there is a differentiation between time spent looking at the text, among the 'distracted' and the 'non-distracted' group, with the 'distracted' one spending less time, a fact that is further supported by Anderson (2009, p.64), which refers to Corbetta and Shulman (2002) experiment, proving the biological differentiation between goal-oriented attention and stimulus-oriented attention. In this case, the goal was to read the text, and the stimulus was the ads.

However, both eye-tracking and results can be questioned. Their potential denial arises from a theoretical and a technical point of view. Elaborating more on this, Aulikki Hyrskykari in her work 'Utilizing eye movements: Overcoming inaccuracy while tracking the focus of attention during reading', 2006, mentions that the point of gaze is highly correlated with visual attention which is a sub-category of focus attention, pointing out '*The focus of visual attention does not necessarily coincide with the focus of attention*.', meaning that although subjects may have spent a certain amount of time looking at the text, that does not necessarily mean that they have paid attention to the text that whole time. But even if it is accepted that the subjects had their focus of attention at the text the whole time because they were in a study environment, making them fully concentrated, there is the accuracy barrier, on which Hyrskykari elaborates, enumerating the three sources of inaccuracy of eye-trackers: *inaccuracy of measurement*, *drift of calibration* and *biological characteristics of an eye* (p.5), and although the two former ones can be encountered, the latter one is a ubiquitous feature.

Continuing, to deduce conclusions about the ability to retain information in the short-term memory, a questionnaire was used. The results showing that the 'distracted' subjects had a lower performance than the 'non-distracted' ones, as expected, are supported by the H1 & H2 which were proven partially true. Furthermore, it is argued that serial bottlenecks in information processing are created when focusing on more than one processes simultaneously (Anderson, 2009 p.63), and thus, making harder for the reader to retain the information, when the text is accompanied by ads. Worth mentioning, is that the rather low average score on both groups, roots in the fact that the ability to keep items in the short-term memory diminishes over a short period of time, proven by Baddeley, Thomson & Buchanan, 1975, in the so named 'word-length experiment' (Anderson, 2009, p.153).

On the other hand, there are causes for which the questionnaire results are not irrefutable. In the study, an important variable has been neglected by not being mentioned at all, that of the depth of processing, a theory argued by Craik and Lockhart, 1972 (Anderson, 2009, p.151), which claims that what is critical, is not the amount of time spent rehearsing the information, but rather the depth to which the information is processed. Although it is difficult to measure this variable, it is indeed of great importance, because many subjects, independently of the group they were part of, might just have grasped through the text without really processing the information they read, resulting in a non-representative deduction. Besides, the incidental vs intentional learning concept, expand the previous argument, stating that '*It does not seem to matter whether people intend to learn the material; what is important is how they process it.*' (Anderson, 2009, p.171). In this case, participants were incidentally learning, meaning that they did not have the intention to memorize the text's information given the fact that they were not aware of the following up questionnaire. Thus, as previously demonstrated by Hyde and Jenkins, 1973, levels of processing are the determinant factor of the amount of information remembered (Anderson, 2009, p.171).

In retrospect, notwithstanding the fact that the research has solidly grounded theoretical background and arguments, that show the results as being supported, there also exist some doubtful and arguable, aspects and variables which must not be disregarded. For that reason, as alluded previously, it is impossible to reach an unquestionable conclusion, but rather a partially correct one.

My answer to question 2

In the attempt to pursue the lines of thought of individuals, through the lens of visual attention, a study was carried out, comparing the ability to remain focused when reading a text, depending on whether the subjects were or not distracted by ads. The outcome proved that subjects which had the text accompanied by ads spent less time looking at the actual text, while their performance on the follow up questionnaire was poorer. The results were based and extracted from the collected data during the experiments, experiments which implementation can be questioned and doubted, mainly from a methodological perspective. Therefore, in this review, a critical attitude will be adopted, trying to both point out the wisely chosen implementations and at the same time detect the ambiguities and potential flaws.

To start with, opting for the ability of retaining information from a given text, in relation to attention level, as a dependent variable, is a way to directly address both the response to the research questions and fulfil the hypothesis, strengthening the credibility of results. This is further supported by the corroborative theoretical grounding regarding the attention and the short-term memory, in question 1 and by the fact that the variable can be manipulated (meaning that each subject can be assigned to one of two experimental groups, where each group represents different types of the independent variable), so that it can be determined to what extent it influences the dependent variable (Bryman, 2012, p.50). Continuing with the independent variables, choosing two of them, distraction on a webpage and no distraction on the other, is a right decision on one hand, as it directly influences the dependent variable (Bryman p.711); examines the desired conditions; and also coincides with the internal validity matter which concerns with '*whether a conclusion that incorporates a causal relationship between two or more variables holds water'* (Bryman, 2012, p.47), i.e. checks if the independent variable is the actual cause of the occurring variation in the dependent one. But it can be reproached for not being clear about the different kinds of distractions that a webpage may have (static, animated, moving) and which of them are being tested, mentioning only '*distracting advertisement-like sections around the text*' (Report paper, p.7). Lastly, coming to the control variables, they have been selected accordingly, as they are meet the criteria of the definition '*variables deemed to have an impact on the relationships between independent and dependent variables'* (Bryman, 2012, p.306).

Analyzing the independent variables, certain material was used or produced to represent them. First, the text was inspired by an actual article in English but generated by ChatGPT (n.d) and retouched by researchers. Although the impartial choice of the text, meaning that the probability that a subject may be an expert in space manufacturing, was extremely low and thus reducing biased experiments, the difficulty of the text is refutable, as it was claimed to be a difficult one, resulting in an even harder information retention task. This is a problem as it weakens the ecological validity of the results, which deals with the question the question of '*whether social scientific findings are applicable to people's every day, natural social settings'* (Bryman, 2012, p.48), that is to say, that most people do not usually read such difficult scientific texts in their everyday lives (obviously that is not a generalization but rather a fact concerning the vast majority). Moreover, speaking of the eye-tracking, even though it is the best alternative to capture the shift of attention ('*a shift of attention often precedes the corresponding eye-movement'*, question 1); this method implies fundamental inaccuracies (also mentioned in question 1), with primary being the biological characteristics of an eye. Adding further to the foregoing point, there is also the technical barrier which cannot be easily overcome, as for materializing the eye-tracking, a supposedly 'free-to-use software', which used the laptop's webcam was utilized, introducing bias to the outcome, beforehand, fact also

mentioned in the report paper (p.18 & p.23). Coming to the questionnaire, although it is not directly related to the independent variable, it implies an indirect connection as it does support the detecting of the amount of information being retained. Speaking of the nature of the questions, they were closed-format questions, because doing so, facilitates the statistical analysis and reduces the variation (Report paper, p.15 – Bryman, p.249-250). Nevertheless, they have certain disadvantages which may have contributed to introducing a bias to the outcome, a bias which might have been minimized if open questions were used instead. To prove this, p.20 of the Report paper, states that '*…leading them to click on the first answer that seemed correct in the questionnaire form, when another answer may have presented itself as 'more correct' would they have read that far*', a behavior which may have not been observed in the open questions case, because then the subject could answer in its own terms without being restricted by the provided answer options (Bryman, 2012, p.247).

Alluding to the results, they supported the initial hypothesis (even though partial support for H1 & H2), as expected. As it was priorly analyzed though, there are not few influencers which may have biased and distorted the results. Regarding, the time spent looking at the text, it was directly affected by the eye-tracker's inaccuracy, enfeebling the measurement validity, argued by Bryman which says that '*the assessment of measurement validity presupposes that a measure is reliable*', and supplementary that '*If a measure is unreliable because it does not give a stable reading of the underlying concept, it cannot be valid, because a valid measure reflects the concept it is supposed to be measuring.*' (2012, p.47). What is more, the questionnaire performance was also impacted by aforementioned factors, permitting thus to question the validity of H3 & H4. Another important factor, which affected overall the research results, is the age group which was restricted within the range 18-30, making the sample constituted only by young adults. It would certainly make a difference selecting subjects within a wider age group as it will be more representative of the population, hence making possible and less incorrect, the generalizations. And with that, the external validity requirements could be met, because it concerns the question of '*whether the results of a study can be generalized beyond the specific research context'* (Bryman, 2012, p.47).

Concluding, although at a first glance the methodology and results may seem plausible, having a more in-depth look would make an individual reconsider some aspects of the study. However, in the end, an answer to the question of how much the end result would then be affected if those mentioned influencers were confronted, can only be deduced if more experiments and similar studies are performed.