Psychology of entertainment
van Driel, K.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
“So tell me, what have you learned during your years of research?” the tutor asked the student.

“Well, that’s easy,” the student answered quickly: “I have learned that in scientific research it is not sufficient to constrain the possible confounding variables. And, another thing, that scientific evidence is much like judicial evidence often at best circumstantial.”

Probably because grand statements like these are what the Dutch call: “grote stappen, snel thuis [big steps, home soon],” the tutor asked for an explanation.

The student then referred to a simple biology experiment he had performed in high school. The experiment, a replication of the famous experiments that inspired Fisher (1935) to develop the analysis of variance technique, addressed the effect of artificial fertilizer on the growth of plants. In order to measure the effect, ten conditions were created, and the plants in every next condition got a little bit more fertilizer. So the independent variable was the amount of fertilizer measured by weight. To measure the dependent variable, the growth of the plant, a ruler with a metric scale for length was used. Every day the plants received a fixed amount of fertilizer and their growth was measured with the ruler. In order to control possible confounding variables, all plants got the same amount of water, were placed on the same radiator to get the same amount of heat and were placed next to each other in the window, so they all got the same amount of sunlight. The same quality of the fertilizer used for each condition was guaranteed by using the same source. If the results had indicated increasing plant length for increasing amounts of fertilizer, most of us, or at least some of us, or at least this student, would have thought this research to be a success. The reality, however, was different: the same plant length was measured every day for all conditions, i.e., irrespective of the amount of fertilizer
used. Of course the student was puzzled, because he refused to believe that fertilizer had no effect on plant growth, as there were too many indications that fertilizer should have a positive effect. Then the student asked himself whether his research design was fully operationalized and which conclusions exactly were permitted by this design.

This experiment is brought up for two reasons: a) because it may typically illustrate the methodological choices that were made concerning the demarcation of the research domain of this thesis, b) in such a way that these choices are comprehensible for a larger audience than the insiders in cognitive psychology and research methodology. At first sight, the experiment does appear to be a classic, but as we already know it may not be, that is, the way the research was performed and presented may be incomplete. Many explanations for the absence of the hypothesized effect are possible, including the explanation that fertilizer has no effect on plant growth. However, the failure to confirm the claim may also lead to a search for a new measuring procedure (Cronbach & Meehl, 1955), as there are often many ways to measure a specific variable. In the case of the fertilizer experiment plant length was selected as an indicator for growth, but now imagine a particular species that grows at a constant speed in length, but with more leaves for increasing amounts of fertilizer. In that case, one could suggest that the weight or volume of the plant would have been better indicators for growth in relation to fertilization, exemplifying that one has to make sure to select the most appropriate operationalization.

In doubt about the most appropriate measure for plant growth, one could use the same experiment as described above to find out. In that case there is a shift in what is called the unit of analysis (Babbie, 1998), instead of focusing on the effect of fertilizer on plant growth, the focus shifts to which is the better indicator for plant growth. Imagine that we conduct the same experiment, measuring both plant length and plant weight. Now, suppose that we find all plants attaining the exact same length, but with increasing leaf density and weight, as the amount of supplemented fertilizer increase. Since we expected fertilizer to have an effect on plant growth, the conclusion that weight is the better indicator in order to observe the expected differences in growth, may be a plausible one, despite the fact that length is still an aspect of growth in this case. Although this thesis consists of four chapters, it comprises six investigations; the first three chapters address one investigation each,
whereas the final chapter discusses three investigations. These four chapters may be classified into two groups by the unit of analysis. The unit of analysis in the first two chapters focuses on measurement methods, as in the fertilizer experiment where weight and length were compared. The final two chapters can be compared to the initial fertilizer experiment, where the effect of fertilizer on plant growth was studied, having psychological models as its units of analysis.

The current thesis aims at the development of a model for the cognitive processing that takes place during the evaluation of cultural products. “How can psychologists claim to measure intelligence when they cannot define it?” the journalist Walter Lippmann (1923) questioned and it may have been this very comment that led Harvard psychologist Edward Boring propose in the 1920s that intelligence is nothing more than what tests of intelligence measure (Sternberg & Grigorenko, 1999). However, as the problems with measuring plant growth in relation to fertilization indicate, this problem is not reserved for psychologists. According to Cronbach (1990) this confusion emerges from the notion that the starting point has to be a definition, while the opposite question would make better sense: “How can one reasonably define a construct until one has made many pertinent observations? Cronbach (1990).” So, in order to define the cognitive processing that takes place during the evaluation of cultural products, one would have to start making observations, but with what measurement method?

Unlike the simple case of plant growth, it is not so obvious which is the most appropriate method to measure cognitive processes. Nevertheless, two prominent ways to measure cognitive processes were considered, the approach taken by De Groot (1946; 1978) who created sequences of cognitive operations using thinking aloud protocols, and the approach discussed by Sudman, Bradburn, & Schwarz (1996) using questionnaires for the evaluation of cognitive processes. Mellenbergh, Kelderman, Stijlen, and Zondag (1979) have introduced a method to use questionnaires as deductive confirmatory instruments, by combining facet theory and confirmatory factor analysis.

The first, exploratory, approach for measuring cognitions is based on the chess player research of De Groot (1946). De Groot created sequences of cognitions to identify different strategies in solving specific chess problems, by asking his subjects to think aloud. Similarly, it is possible to create sequences of cognitions
taking part in the evaluation of a cultural product. This set of sequences may subsequently be clustered in order to identify different types of problem solving strategies or different types of cognitive evaluations of cultural products. The second, confirmatory, approach uses questionnaires to validate an a priori theory.

Similar to the latter fertilizer experiment, the unit of analysis was the preferred method. One would be comparing apples and oranges when comparing both approaches in one experiment, because of the extreme difference between the two: exploratory versus deductive. Therefore, two separate investigations were conducted to evaluate the merits of both approaches, using accepted theories to control as many variables as possible. Based on these evaluations it would be decided which strategy would guarantee the most promising approach for the development of a model for the cognitive processing of cultural products.

Chapter 1 discusses the evaluation of the first method. This method is an agreement measure designed to obtain similarities between sequences, such as sequences of cognitive abilities. The similarities may be used for further analysis with, for instance, cluster analysis in order to obtain a taxonomy of cognitive strategies for the evaluation of cultural products. The chapter discusses an evaluation of the method that was found mathematically flawed. Suppose that plant length in the fertilizer experiment would have been measured with an elastic band instead of a ruler of solid fabric, it could have been argued that the elasticity of the instrument may have been the reason that plant length as measured failed to discriminate plant growth. Similarly, it is proven that the method discussed in chapter 1 is unable to discriminate sequences of cognitive abilities.

Chapter 2 is an evaluation of the confirmatory facet design questionnaire approach. The triarchic theory of human intelligence (Sternberg, 1985), which is widely accepted by psychologists, is used as content for the questionnaire. Unit of analysis, at this stage, is the research strategy that uses a facet theory (Guttman, 1954, Canter, 1985) to design a questionnaire with typical performance measures (Ackerman, 1994), i.e., self reports. As opposed to maximum performance measures such as problems or tasks that have to be solved, one could argue that self reports introduce a biased source of variance, for instance, because individuals are not able to make a pure estimation of their capacities, but also consider their background or development.
This criticism may be illustrated by the controlled amount of water in the fertilizer experiment, for is it possible to study the effect of fertilizer by controlling the amount of water given to each plant? After all, plants that grow faster need more water and by limiting the amount of water we may never find out the full effect of the fertilizer. Furthermore, too little water may turn high concentrations of fertilizer into poison for the plants, and too much water for the no fertilizing condition may drown the plants. Farmers in the droughts of Australia know this problem, for it is no use fertilizing when no rain is expected. Similarly, one could ask whether it is possible to study intelligence independent of one’s upbringing and environment? Good breeding may lead to an overestimation of the cognitive capacities of an individual compared to less fortunate peers, but the other way around, it could also be possible that someone with lower capacities is ‘poisoned’ by a stimulating environment. Consequently, increased quality of the environment may also have an inhibiting effect on the full development of the capacities of certain individuals. This would be an important issue if the units of analyses were the individuals and the demarcated construct of interest were cognitive capacities as defined by a predisposed capacity irrespective of nurture. However, in the present case, the research objective is not making accurate estimations of individual capacities but to explore whether the method may be a useful application to uncover cognitive processes.

As the first approach lacked mathematical justification, and the second supplied a successful operationalization of a cognitive process theory of intelligence, the latter approach was chosen as the most promising for a successful development and evaluation of a model for the cognitive processing of cultural products. As the investigation established the research strategy, the unit of analysis could safely be shifted to the model for the cognitive processing of cultural products.

In chapter 2 another possible unit of analysis, in addition to the method and the theory, has been introduced: the individuals. In the studies throughout this thesis the individuals are never the unit of analysis, however, they are being observed and may be referred to as units of observation. By observing individuals, first an evaluation of methods is made, and second, content models are evaluated. Chapter 3 discusses the pilot study, introducing a preliminary model for the cognitive processing of literary works. Now another source of variance enters the research: the
Psychology of Entertainment

literary works that also may be regarded as units of observation.

In this chapter, and in the next, the literary and cinematographic works are not always manipulated, i.e., not all individuals read or watch the same work, but the subjects select the works. In terms of the fertilizer experiment, this is illustrated with an experimental design in which all conditions had had access to an unlimited source of water from which they could select whatever amount they needed. Of course, one could argue that an experimental design with varying levels of fertilizer and varying amounts of water, studying the interaction between the two, would have provided a better design, but one should not forget the advantage of hindsight. Suppose that the current experimental design had demonstrated the hypothesized effect of fertilizer on plant growth; in that case such a costly experimental design would have been an over-operationalization of the hypothesis that fertilizer affects plant growth. Similarly, it was considered at this stage in the research program preliminary to control the literary works, for the aim was at establishing the presence of the hypothesized cognitive processes. However, because self-selection of literary works by the subjects could have a possible confounding influence, relevant subject characteristics were taken into account, such as personality traits and cognitive abilities.

Let’s go back to the question whether the fertilizer experiment was sufficiently operationalized. Both the dependent and the independent variable were operationalized with measurement units that can be connected to accepted measurement theories, i.e., length as measured by a metric ruler and weight as measured by calibrated weighing scales. A thermometer, another widely accepted instrument, even measured the heating of the soil. However, the amount of sunlight was not measured for every plant; but it was made plausible that all plants got approximately the same amount of sunlight by placing them in similar positions next to each other behind the same window. Similarly, the quality of the fertilizer was not measured either. However by using the same source, i.e., the same bag of fertilizer, it was made plausible that weight sufficed to measure amounts of fertilizer, since all the fertilizer used had the same quality. This illustrates the condition that not all variables necessarily need to be controlled by measuring them; sometimes it may suffice to make plausible that certain variables are unlikely to have an influence under certain conditions.
Where chapter 3 is an account of the pilot investigation of the model, chapter 4 forms a report of the final research, addressing various aspects of validity: construct, internal, external, and predictive validity. These terms will be clarified for the novel re-creative readers of this preface. Construct validity, referring to whether the constructs of theoretical interest have successfully been operationalized, has already been addressed, e.g., the inadequacy of plant length as an indicator for plant growth in relation to fertilizer for these specific plants. In chapters 3 and 4 questions regarding the construct validity address the question whether the hypothesized cognitive processes could be discerned. Internal validity refers to the extent to which conclusions can be reached about the effect of the independent variable on the dependent variable. Internal validity can be exemplified by the question raised in the fertilizer experiment whether fertilizer has an effect on plant growth. In this thesis the relation between levels of cognitive activity and aspects of the literary and cinematographic works are investigated. The external validity addresses the generalizability from the research setting beyond the boundaries of the performed study, e.g., does fertilizer only have an effect on plant growth in leaf density for all species of plants, or are there species of plants that would benefit from growth in height by fertilizer? Is the model for literary works also applicable to cinematographic works? Predictive validity, finally, demands as its name suggest that for instance plant weight can be predicted from the amounts of fertilizer used, or in the present case whether the appreciation of motion pictures can be predicted by cognitive activity.

“Well, that’s it,” the student said.

“Now, it is still not clear to me what have you learned from this research. You already seemed to know everything when you performed the fertilizer experiment in high school.”

“Do you need to ask,” the student replied, “Not only did I then learn the effect of fertilizer on plant growth, but now I gained insight on the cognitive processing of cultural products. Furthermore, back in high school I was under the impression that the research was a failure because I controlled all the variables. I thought that my measurement was accurate, the chosen instruments the only ones possible, the method valid in all aspects of validity and I ignored the possibility, of
Psychology of Entertainment

an interaction effect between water and fertilizer. Now I know, that a shift in the unit of analysis may put unexpected results in a different perspective."
REFERENCES


