Mental Content-based Analysis of Transfer

Sacha Helfenstein and Pertti Saariluoma[1]

Department of Computer Science and Information Systems

P.O. Box 35

FIN-40014 University of Jyväskylä

Finland

E-mail: sh@cc.jyu.fi, psa@cc.jyu.fi

[1] corresponding author

Abstract

The reported study examined the role of mental contents in human apperception and reasoning. Based on Duncker's (1935) classic tumour task and the transfer setting used by Gick and Holyoak (1980, 1983) we devised an empirical paradigm that allowed us to investigate the influence of three different thought models of rays confluence (*additive*, *balancing*, and *distribution-based*) and two distinct types of spatial images of rays (*compact* vs. *diverging*) on people's problem representation and judgments. In each of three consecutive experiments, a different sample of subjects, divided into two ray image conditions, was primed with thought model-consistent learning examples and pictorial illustrations of rays. Although all subjects performed the same task and were required to reason about the same confluence schema, they differed significantly in their judgment of confluence effects applied to the tumour task. These differences reflected the contrasts between the thought models and the ray image contents and substantiate our claim that a schema-based analysis of transfer, and reasoning in general, is alone not sufficient enough to explain interindividual and intercontextual differences that are based on distinctive mental contents in reasoners' apperception.

1 Introduction

1.1 Content-based psychology and transfer research

It is a very natural assumption that information content in mental representations can explain human behaviour. Approaches that use mental contents as the explanatory framework for investigating psychological phenomena can be called content-based (Saariluoma, 2002a, b, 2003). We have been developing the foundations for this type of psychology for over a decade and have applied it to various research questions that can best be solved by employing a content-based approach (Saariluoma, 1990, 1992, 1995, 1997, 2001, 2002, 2003; Saariluoma & Hohlfeld, 1994; Saariluoma and Kalakoski, 1997, 1998; Saariluoma & Maarttola 2001, in press).

One of the key issues in content-based psychology is the analysis of the apperceptive process and with that the distinction between perception and apperception (e.g. Saariluoma, 1990, 1992, 1995, 2001). When putting the focus on the contents of human mental representations, it becomes necessary to understand that many of them do not have their origin in currently perceivable physical stimuli. In fact, a large part of what is represented in our mind consists of non-perceivable content. Contents that cannot, or not directly, be perceived, e.g. tomorrow, infinity, possible, constitution, eternity, are incessantly incorporated in our mental representations. For this reason alone, it is important to make a difference between perception and apperception (see, for example, Kant, 1781; Stout, 1890; Wundt, 1920). It is important to ascertain that our mental representations are to a great extent of a conceptual nature – and not solely perception-dependent; and even less stimulus confined.

Apperception can simply be defined as the mental process that selects and constructs the information contents in our mental representations. These representations can refer to the external world as much as to our inner life. Furthermore, apperception unifies perceptual information contents, i.e. information encoded from physically present stimulus environments with non-perceivable kinds into mental contents, which controls our ongoing behaviour and mental activities, e.g., the flow of our thoughts. In order to understand apperception, we have to be able to break the contents of the mental representations into content elements and investigate how they have been unified and how they can explain observable behaviour. This deconstruction challenge is not one of atomizing our mental life, but a search for meaningful units of mental constituents and their relationships to each other in our mental representations.

So far, we have been able to illustrate that there are mental contents that effectively explain human thought errors and economy of thinking. In chess, and also in architectural design, it can be shown that people use functional rules, often subconsciously, to integrate content elements. These rules are essentially reasons, i.e., they tell why some elements belong to a mental representation (Saariluoma, 1990, 1992, 1995; Saariluoma & Hohlfeld 1994; Saariluoma & Maarttola, in press). Sometimes, the functional rules are incorrect, i.e. "pseudo rules", and lead into partially false representations (Saariluoma, 1992, 2003; Saariluoma and Maarttola, in press).

Another interesting property of mental contents is apparent in the so called thought models. Thought models are a type of mental content, which are used to guide the human thought process. These models can sometimes be very risky. For example, a model called Chernobyl entails the idea that it is not necessary to follow safety regulations. This model has been termed Chernobyl as more than 60% of work accidents in Finland follow from work practices that are generally known as being dangerous (Saariluoma, 2002, 2003). Thus, risky thought models can be used to explain faulty thinking in working life.

In this paper, we focus experimentally on the role of thought models that are applied to a classic transfer problem in order to gain additional information about the nature of apperceptive. We are especially interested in image-based spatial information. Transfer task environments are ideal for content-based investigations and apperception research. According to the classic experimental paradigm for the study of transfer tasks, definable mental contents are activated in the minds of the subjects during the initial learning situation. Subsequently, the function of these activated contents for encoding, representation, and solving of the transfer task can be effectively analysed.

A second and equally important reason to choose the transfer phenomenon is the solid tradition of transfer research. This allows us to design the experimental tasks in a much less intuitive manner as in the case of using less investigated tasks. The work of Duncker (1935) and its incorporation into transfer research by Gick and Holyoak (1980, 1983; see also Keane, 1987) has uncovered a number of important properties of transfer processes in thinking. These investigations effectively support our attempts to use transfer for the investigation of apperception.

Gick and Holyoak (1980, 1983) used a problem solving task developed by Duncker in the 1930s (the so-called "tumour problem") and devised an analogical story for this problem (the so-called "General story"), which was given to the subjects as a learning task (see Appendix 4.1, for both story lines). They found an effective positive transfer, which was explained in terms of a schema theory of problem analogy. Here, we continue this line of research. However, we wish to demonstrate that two schematically invariant solutions to a problem can eventually lead to diverse mental representations. The reason for this lies in the mental contents of our representations (or thought models, as is studied in this paper), which may differ despite commonalities on a schematic level. In extreme cases, mental contents may make a difference between a positive and negative transfer.

For this reason we shall argue that it is not sufficient to explain the transfer phenomenon on the ground of schemata alone, but it is essential to base an explanation on the analysis of the information contents of the schemata. Arguing in such terms that probe the foundations of theoretical development, it appears valid to state that the mere concept of *schema* provides a scientific language that is not sufficiently powerful to express and explain the full range of transfer results (what is meant by *the power of expression* is explicated in Saariluoma, 1997). Specifically it is not powerful enough to profoundly explain the high frequency of findings concerning the failure of transfer, typical of analogical problem solving and schema-based research. Hence, this paper intends to provide further support for our general suggestion that content-based explaining is essential in investigating a number of questions in human behaviour.

Content-based explaining has an additional advantage. It allows for the integration of element-based and schema-based transfer. Transfer is caused by the contents of representations, which includes different types of mental constituents – elements and schemata - and therefore makes a fundamental distinction between them unnecessary. The decisive issue is the contents.

1.2 General empirical considerations

The problem we are dealing with here is the one of experimentally induced transfer. That means we, as researchers, choose and design learning and transfer tasks based on our view of similarity. In this respect our investigation shares its basic empirical paradigm with the one employed in previous transfer research, on which the current paper draws. The main difference to the current study lies with the explanatory concepts that govern the choice of experimental variables and the interpretation of the gained results.

Gick and Holyoak (1980, 1983), for example, constructed their experiment on the theoretical belief that transfer is a function of the degree of schematic congruence between two problem solutions (or relational correspondences between attributes across domains). As indicated, analogical transfer research has been haunted by findings demonstrating the absence of transfer gains or even a negative outcome (see, for example, Campione, Brown & Ferrara, 1982; Detterman, 1993; Reed, 1987; Simon & Reed, 1976). This has not been much different from the research of Gick and Holyoak, where the spontaneous use of their story analogy to come up with a transfer solution to the tumour problem was present but small. One could say that the results lagged somewhat behind the predictions based on the perspective of the analogical or schematic transfer theories. Less than one third of their subjects appeared to be able to spontaneously and successfully transfer the solution schema from the General story (as well as other story analogies) to the solving of the tumour problem. Non-spontaneous conditions are not considered, because transfer due to the presence of a hint, or as a consequence of an established routine to look for analogies, is quite a distinct issue in everyday learning. It is rather an issue of application than of carry-over in its full sense: Everybody is in essence able to find a needle in a haystack provided we tell him or her that it is there. The low findings for spontaneous analogical or schematic transfer across domains are usually explained with superficial differences and the “barrier”-effects they exhibit in our reasoning.

Other transfer researchers, such as Anderson (1983, 1985, 1993), developed a distinct approach to transfer that focuses on elementary mental communalities between tasks: the production rules (see also Moran, 1983; Polson & Kieras, 1985; Singley & Anderson, 1985, 1989). Thorndike and Woodworth (1901a, b, c) already laid the foundations for this view at the beginning of the last century. It is clear that this transfer paradigm too, delivered some noticeable results but it has its own difficulties especially with regard to the account of negative types of transfer (Singley & Anderson, 1989).

As laid out in the introduction, here we intend to support our case concerning the theoretical and empirical value of a different kind of mental constituent. We propagate *mental contents* and investigate their role for thinking and behaviour in general and transfer specifically. In general, our argument is also related to Novick's conclusions about the representational predicament of transfer (Issing, Hannemann, & Haack, 1989; Novick, 1990) but it builds on a distinct view of how representations are constructed and what the role of the apperceptional processes involved in their construction for transfer is.

2 Experimentation

2.1 Prior qualitative investigation

Although the demonstration of positive transfer is the main road for developing and establishing a theory of transfer, it is well advised to complement it with the examination of transfer failures. By examining transfer failures we are able to uncover blind spots in theories instead of simply blaming the experimental design when the empirical findings fall short of the theoretically based predictions.

Prior to the reported experiments we conducted a small study based on qualitative interviews, which was constructed to serve three main intentions: The first purpose was to confirm the findings of Gick and Holyoak in a no-hint learning context, using their original General story analogy. The second purpose was to rule out that the poor transfer findings were limited to the source analogs devised by Gick and Holyoak. For this an alternative source analog was constructed. Third and probably the chief objective was to prompt the subjects to elaborate on the possible transfer sources for their solutions, which differed from the transfer solutions envisaged by Gick and Holyoak. We encouraged the subjects to give their own account of why they did not come up with the General-tumour analogy themselves. For this, the schematic analogy was made so salient to them that they experienced the „Aha"-effect and thus recognized the similarity.

Eight subjects read one of the devised learning stories („tumour story"-analogies) in between two others, here irrelevant stories, in the context of a memory recall experiment and were then confronted with the tumour story in a separate experiment conducted immediately after the first one. They were then asked to provide a deliberate amount of solutions to the tumour problem during a five minute time period, and then had to point out and draw their own favourite one. The subjects were urged to talk-aloud during their problem solving. The talk aloud protocols were recorded for analytical purposes.

Several important findings emerged from this qualitative study. Firstly, poor spontaneous transfer of the General solution (as well as the own analogy) to the tumour problem was confirmed for all learning stories. Only one subject spontaneously came up with analogy solution. However, all subjects provided reasonable solutions to the tumour problem and mostly were also able to give a personal transfer source for these solutions. All subjects also experienced the „Aha"-effect and were able to elaborate the analogy without further help from the experimenter, once the link was pointed out to them. This indicates that the findings of poor spontaneous transfer may be less a problem of the lack of transfer or spontaneity and more one of the inappropriateness of the presumptions made about the mental relationship between the learning and the transfer situation.

Finally, it became clear that the main obstacles for General- to tumour story transfer might be largely inbuilt into the tumour problem itself. The subjects argued to know too little about the properties of radiation, and when they commented on their intuitive and naive representations of rays it became clear that these may be irreconcilable with the type of transfer intended in the Gick and Holyoak task. With regards to the solution of the General story, some subjects also mentioned their incertitude about the achievability to effectively control and apply the armed forces' full potential when dispersed and converged. These latter issues shall be further elaborated in order to make the general rational of our experiments better understandable.

2.2 General rationale

Our research builds on well-founded theoretical considerations and on conclusions drawn from our qualitative observations of people attempting to solve the General-tumour story analogy used by Gick and Holyoak (1980, 1983). The basic argument reads that transfer explanations based on the schematic assumptions of the General-tumor story analogy presupposes the availability (or existence) of ideal mental contents in how people represent rays and their confluence. What we need to do is to identify these contents and to show that in absence of these ideal apperceptional contents or in the case of fundamental differences from these ideal representations of radiation, schematic transfer from the General to the tumour problem may lack meaning and fail.

The roles of two implicit assumptions concerning the contents of problem representations are investigated. The first one deals with the circumstance that in the case of the General-tumor problem not only the critical solution schema must be transferred, but for this to succeed the mental image of radiation needs to incorporate a specific *content*; itself probably transferred from earlier learning experiences concerning radiation. In concrete to transfer the *dispersion-convergence* solution from the General story to the tumour problem the medical rays need best to be imagined in a compact, laser-like fashion (see Figure 1). This is obvious from the original version of the tumour problem (Duncker, 1935, p. 2).

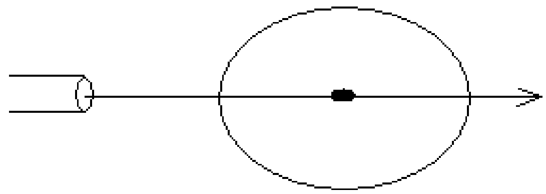


Figure 1: Illustration Duncker (1935) provided his subjects with while solving the tumour problem

Only if the rays are imagined in this fashion, their convergence (or better *confluence*, as it shall be referred to during the course of this paper) at a single point really makes sense. From the interview protocols it became clear, however, that the subjects imagined rays in a rather chaotic, diffuse, and in most cases diverging manner.

We also assessed this same issue in a little questionnaire administered to 15 subjects in the context of a lecture at the Chydenius Institute in Kokkola, Finland.

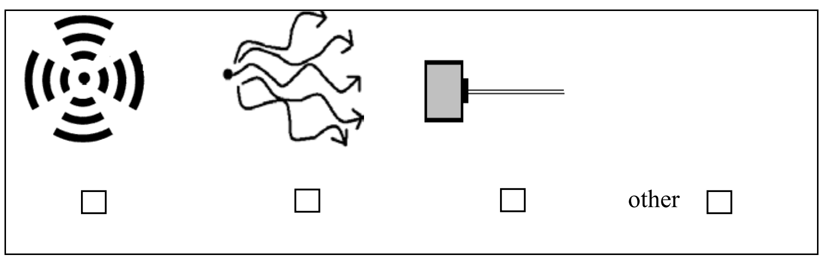


Figure 2: Options for ray representation (image)

The subjects were required to explicitly point out their imagination of radiation by choosing from one of three given representational options or provide their own (see Figure 2). Here we found that only *one* subject crossed the compact, laser-type of radiation image, which was originally proposed by Duncker as the ideal type of ray-representation to solve the tumour problem. In contrast eleven subjects reported to imagine radiation in the diverging fashion of the first pictorial on the left and two in the more diffuse fashion depicted in the second image. One subject provided a personal pictorial account of his or her imagination. It must be noticed that the question about ray representation was preceded

> by putting the subjects into the context of using rays for medical purposes for destroying tumours, and can therefore not be purely seen, e.g., as an effect of the well-known radiation symbol. This representational factor concerning the image of rays may critically influence the way subjects' judge the effectiveness and harmfulness of the ray confluence technique when attempting to destroy a tumour.

A second implicit assumption made in the General-tumour story analogy concerns the *thought model* involved in representing the confluence effects. This applies to the dispersion and confluence of the armed forces (in the General story) as well as to the knowledge of the (physical) properties of rays with respect to their division and confluence at a point (the tumour). From our qualitative interviews it became clear that the subjects were generally very uncertain about the consequences of confluence.

Here, we were mainly interested in the subjects' representation of ray confluence. Gick and Holyoak's (1980, 1983), as well as Duncker's (1935) assumption was that the subjects understand the effectiveness of a rays confluence on destroying the tumour in an additive way. This means, the energy of four smaller rays converged on a single point equals the energy of one (four times) more powerful ray. At the same time, it is ascertained that the harmfulness to the surrounding tissue would remain unchanged (non-additive). The contradictory potential of these assumptions needed to be investigated.

Since the additive thought model plays a crucial part when judging the rays' effectiveness for destroying the tumour, we designed our first experiment to see how well our subjects applied this way of thinking about confluence to the tumour problem. In Experiment 2 and 3, we extended our approach to evaluate alternative thought models that may influence the way subjects understand and interpret the effect of confluence.

Figure 3 makes the underlying idea of the current study apparent. In the three experiments we assessed consecutively 1) the additive thought model of confluence effects, expressed by a summation of a single input; 2) the balancing thought model, expressed by an average over a single input; 3) distribution-based thought model, expressed by the idea that effect of confluence exceeds the sum of the single input.

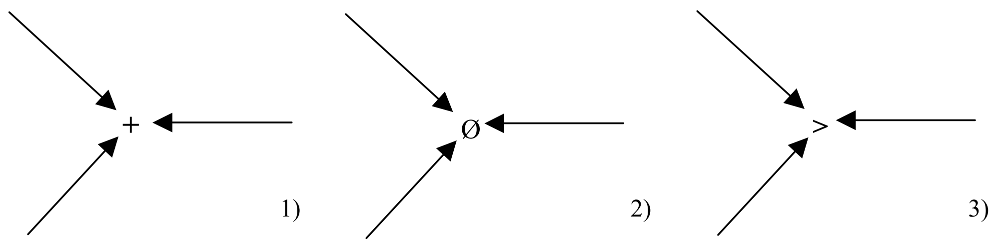


Figure 3: Three distinct thought models involved in the mental representation of confluence effects

In each experiment we included the same basic elements involved in the cancer radiation treatment task (the tumour and the rays) and one and same schema of how they are applied (confluence), but we manipulated the thought model that governs the process of constructing a mental representation of the situation as a whole, and thereby affecting the reasoning.

2.3 Experiment 1

2.3.1 Aim

Experiment 1 had the objective to take a first look at how the subjects understand the relationship between the effectiveness and harmfulness of one big ray compared to four smaller, confluencing rays. In other words, how do subjects think of the effect of confluence in the case of radiation. What we had in mind was that depending on the subjectively available image of radiation and the currently activated mental content (or understanding) of confluence, the same confluence schema applied to the tumour task may result in very different judgments of this technique.

Concerning the diversity of images of a ray (diverging vs. compact; compare first and third options in Figure 2) we were interested in whether the subjects rate the harm-diminishing effect of dividing one big ray into confluencing four smaller rays as less pronounced when the rays are suggested as diverging in nature, compared to a compact image of rays. We assumed the subjects primed with an additive model not only believe that four confluencing smaller rays increase the harmfulness compared to just one small ray, but that this increase is larger for the condition where rays are seen as diverging in nature, compared to an image of a compact ray. The latter part of this hypothesis was derived from the reports of our qualitative interviews, where subjects often referred to the chaotic way in which rays harm the surrounding tissue.

Our second and main objective was to influence and investigate the consistency of our subjects' opinions concerning the additive thought model of radiation confluence. We wanted to make the additive model of confluence even more salient than has been possible by reading the General story, which was not fully convincing in this respect to many of our subjects that participated in the qualitative pre-study. However, based on our observations from our qualitative interviews, we still did not believe that the subjects' judgment of the rays confluence effects would be strictly in line with the ideal additive thought model. If proved right, this would challenge the assumption about the availability of a transfer-appropriate thought model for representing ray confluence.

2.3.2 Method

The experiment was conducted by use of a questionnaire containing two learning examples, a learning task, and a transfer judgment problem. We devised a classical transfer paradigm to activate the desired additive content or principle of confluence during learning and then measured its effect in a transfer problem. We gave the subjects the opportunity to become familiar with the additive model by studying three examples from outside the domain of radiation (see Appendix 4.2 for English translations of the questionnaires).

In the introduction (learning section of the questionnaire) we explained to the subjects that one important effect of confluence is to augment forces in an additive fashion. We provided them with two simple examples and let them judge the third one themselves, as to get familiar with the judgment questioning method. In the transfer task the subjects were urged to judge effectiveness and harmfulness of confluencing (as well as single) rays for destroying a tumour and preserving the surrounding healthy tissue likewise. Thus, the radiation task followed in essence as a fourth example of confluence. It should be pointed out that the subjects were not told that the confluence of rays obeys the same additive law as their learning examples did. We just wanted to make sure that the subjects understood the additive principle by completing the learning section and were potentially able to integrate this mental content into any other representation of a confluence situation. The questionnaire used a relative magnitude estimation technique, in detail described by Stevens (1961) to measure stimuli perception. The subjects could freely assign any anchor value they wanted for the effectiveness and harmfulness of one ray (e.g. 1, ¼, 4, 100), and then provide a relational value to describe the effectiveness and harmfulness of four confluencing rays.

The appropriate ray representation was suggested by the use of two different ways in which the rays were pictorially presented in the transfer task. Accordingly, the questionnaires were distributed in two different versions. One with the rays displayed in a diverging manner, the other with compact ray images (see Figure 4).

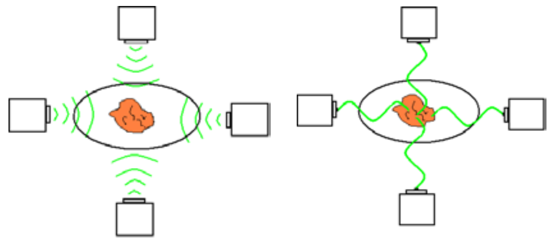


Figure 4: Ray images used in the two questionnaire versions (left: diverging; right: compact)

Nothing in the text explicitly underlined this distinction and it was thus of a rather suggestive nature. Depending on the questionnaire version the subjects could then be divided into two ray image condition groups: *diverging* and *compact*.

2.3.3 Subjects

In all 36 subjects filled in the questionnaire. All were recruited by use of an e-mail list and were enrolled as degree students at the University of Jyväskylä, Finland.

2.3.4 Results

Although the variety of judgments about the rays' effectiveness and harmfulness was our key focus, we excluded those subjects from the analysis whose answer violated some absolute basic logic. This was true for five subjects' judgments, where one small ray was noted as being more harmful than four of the same small rays. It is suspected that these subjects misplaced their answers in the fields on the questionnaire.

The frequency table of the judgments of the remaining 31 subjects shows that, while the majority of the subject judges one small ray to be four times less effective and harmful (relation = .25) than four of the same small rays, a number of subjects view the single ray situation as equally effective and harmful as the confluence situation (relation = 1) (see Figure 5).

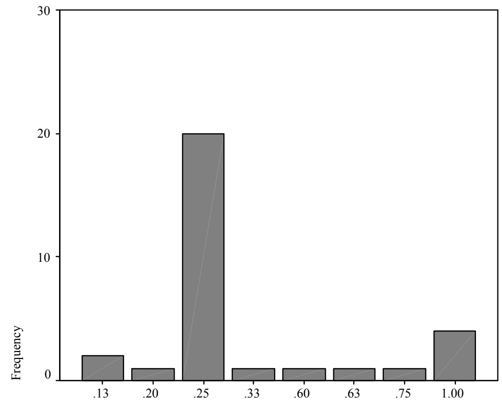


Figure 5: Effectiveness relationship: One small ray compared to 4 converging small rays

For the effectiveness of the judgment of the mean lay at m=.38 (s=.28) for the harmfulness at m=.37 (s=.28). Thus, while the majority of the subjects judged the effectiveness of the confluence schema of administering rays in line with the additive thought model, there was on average a significant deviation from the ideal value of .25 (t(*df*=30)=2.61, p<.05).

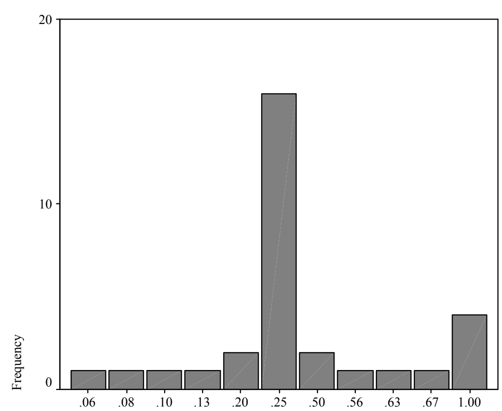


Figure 6: Harmfulness relationship: One small ray compared to 4 converging small rays

Further, the majority of the subjects also judged harmfulness according to the additive thought model. This violates the presumptions inbuilt into the solution of the Duncker task. Only 4 subjects (13%) judged four confluencing rays as equally harmful as one ray alone.

This also means that our prediction about the general apperception of four confluencing rays as more harmful than one proofed to be right. We then wanted to explore further the harmfulness ratios depending on whether the subjects received a questionnaire about radiation being displayed in a diverging manner (n=16) or in a compact manner (n=15).

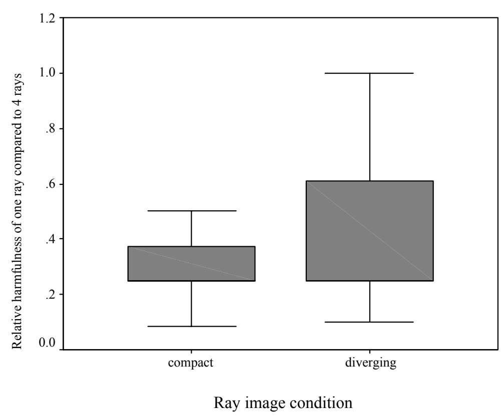


Figure 7: Ratio of harmfulness judgment ($^{1\ ray}/_{4\ rays}$) for diverging and compact ray image condition

As Figure 7 shows, this contrast did not produce the expected picture. The subjects in the diverging ray condition judged the increase in harm, when using four rays in a confluencing manner, in average roughly the same as did the subjects in the compact ray condition ($t(29)=-.93$, $p_{(1\text{-tail})}>.1$). However, as Levene's test for equality of variances shows, the subjects' judgments in the diverging ray condition were also more heterogeneous ($F_{Levene}=2.39$, $p=.133$). That means that more individuals were of the opinion that the increase in harm, when adding rays, was less severe than is suggested by the additive model

2.3.5 Discussion

To summarize, we found that on average the subjects judged four confluencing, small ray as more effective but also more harmful than one single small ray. The latter differs from the ideal understanding presupposed in the tumour task and it is suggested that one reason for this lies with the prior activation of the additive thought model when considering confluence effects. Considering that the rays' harmfulness to the surrounding tissue is the main constraint of the tumour problem, it is a highly relevant finding that only 13% of our subjects would be able to appreciate the functional meaningfulness of the confluence solution for diminishing harm.

In the questionnaire we explicitly stated that ray intensities might be varied in a purely proportional manner (i.e. that one can produce rays of intensity x as well as $n \cdot x$ and $^{x}/_{n}$). Using this computational transformation we may derive from our data that, on average, subjects judged one big ray (i.e. four times more intensive ray as the above mentioned „small rays") to be slightly more harmful and interestingly also more effective than four confluencing smaller rays. That means that, in spite of the learning examples, the additive model was not the only thought model applied to the judgment of ray confluence effectiveness.

The subjects' judgments clearly differed among each other and this interindividual difference was more pronounced for subjects who received the diverging ray image. The different pictorial presentation of rays used in the questionnaire (diverging vs. compact) did not, however, produce the expected effects on the harmfulness judgment. It is unclear whether our purely suggestive technique was not effective enough, whether the additive confluence examples we provided produced an over-homogenization of how people think about radiation confluence, or whether our research assumption was inappropriate. It may in fact also be that the subjects interpret diverging rays as losing energy on the way to the tumour, which does not harm the surrounding tissue, but simply disperses into „space".

2.4 Experiment 2

2.4.1 Aim

The results of Experiment 1 showed that the additive thought model was applied rather well when judging the effects of using a confluence schema for destroying a tumour with rays. It is important to emphasise, however, that the judgments were not univocal. Both findings, and especially the one regarding the application of the additive model to harmfulness challenges the implicit assumptions about people's understanding of radiation confluence necessary to solve Duncker's (1935) tumour task. They also open interesting new content-based insights into the proposed analogy of the tumour problem to Gick and Holyoak's (1980, 1983) General story.

In the current experiment we wanted to make an investigative step towards a better understanding of the discovered variety of effectiveness and harmfulness judgments, which was found in spite of the salience of the additive thought model represented by the learning examples. It becomes clear from Figures 5 and 6 that the second largest frequency was registered for a type of judgment seeing four confluencing rays as equally effective and harmful as one ray (N=4, for both criteria).

Thus, the rationale of Experiment 2 was to establish whether this type of judgment could be represented by a distinct thought model (mental content) applied to the same schema of confluence. Hence, we needed to find such examples of confluence that are in line with that principle of judgment.

2.4.2 Method

For comparability reasons we chose the same procedures as in Experiment 1. We changed the questionnaire from Experiment 1 only where needed to emphasise the new thought model (see under Appendix 4.2). The introductory text was altered so that it stated that an important property of confluence of different input sources was the maintenance of a certain threshold. For simplicity reasons we used the new model in a way that multiple input sources or forces produced the same effect as was reached by having only one input source. For example, mixing liquids of identical concentration of some critical substance leaves the effective concentration unchanged, i.e. no matter how many bottles of Coca Cola I fill into a bowl, the proportional content of sugar remains the same. The same is the case when having only one compared to two water taps providing 30 degrees Celsius of warm water to a bath tub; provided, of course, that the bath tub is well isolated. In Experiment 2 the concentration task was used in the learning task, which the subjects had to judge for themselves. Two temperature confluence situations were described as introductory learning examples before that. It is perhaps worth noting that in reality the model description is not fully adequate regarding the examples provided. Mathematically more correct would be to call it a *balancing* thought model of confluence, since the combination of different input results in a weighted mean of the input variables. The concept of balance shall therefore be used in this paper to describe the essence of the current thought model. However, since all our learning examples as well as the ray judgment task dealt with equal input forces, it proved sound enough to stay with the simple model description in the questionnaire.

The main hypothesis to be tested was a direct offspring of the observation in Experiment 1, namely that a number of people view one small ray as equally effective or harmful as four small rays. We expected that a greater portion of people judge four confluencing small rays as equally effective and harmful as one small ray, and that the average judgment ratio (i.e. the relation of the judged effectiveness and harmfulness of one small ray compared to four small rays) obtained in the current experiment is significantly bigger than the same ratio judged by subjects primed with the additive confluence examples.

2.4.3 Subjects

In all 36 subjects filled in the questionnaire. All were recruited by use of an e-mail list and were enrolled as degree students at the University of Jyväskylä, Finland.

2.4.4 Results

Looking at the central tendencies of the answers given by the 30 subjects that have correctly completed the questionnaire (like in Experiment 1, five subjects were excluded from the analysis due to suspicion of having misinterpreted or misapplied the judgment task), we found a slight increase in the average ratio given compared to the results in Experiment 1: effectiveness: m=.42, s=.28; harmfulness=.54, s=.36. For the judgment of the effectiveness to destroy the tumour our subjects' opinions about ray confluence moved thus further away from the ideal value of .25 (additive model). This deviation, and the fact that it is slightly more pronounced in Experiment 2 are in line with our research assumptions ($t(29)=3.26$, $p<.01$).

For the harmfulness judgments, the observed mean increase points in the right direction, however, remains clearly below the ratio of one, ideally suggested by the threshold maintenance (or balancing) principle. T-tests of the mean differences in the judgments between Experiment 1 and 2 the results for effectiveness as statistically significant ($t(59)=-.55$, $p_{(1\text{-tail})}=.29$), but do so for harmfulness ($t(59)=-1.97$, $p_{(1\text{-tail})}<.05$); while homogeneity of variances in the latter case cannot be assumed ($F_{Levene}=9.855$, $p<.01$).

Looking at the frequencies of the judgments (Figures 8 and 9) we find that answers which rate the effect of one small ray as less than one fourth (.25) of four confluencing small rays have almost completely disappeared.

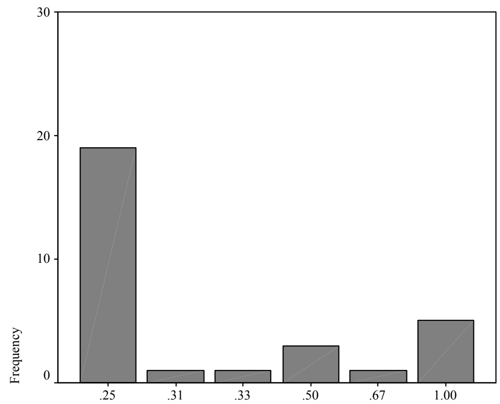


Figure 8: Effectiveness relationship: One small ray compared to 4 converging small rays

We also find that for harmfulness there is a noticeable increase in answers that are in line with the balancing thought model (Figure 9), which was not found as such for the judgment of effectiveness.

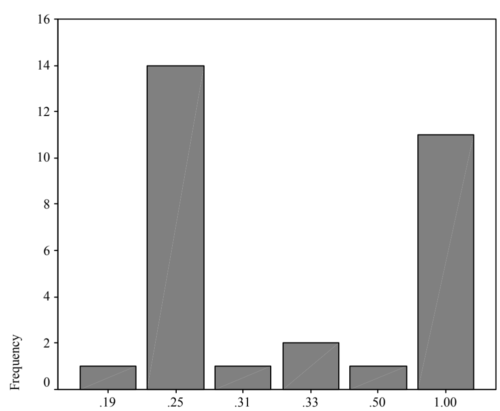


Figure 9: Harmfulness relationship: One small ray compared to 4 converging small rays

2.4.5 Discussion

The results from our first two experiments demonstrated that for the judgment of ray confluence effectiveness with respect to destroying a tumour, it makes a moderate difference whether we activate the mental content the additive model or the one of balance. That could mean that this type of thinking about radiation confluence is rather robust, although not fully in line with the ideal assumptions. However, the same confluence schema, judged with respect to the rays' harmfulness to the surrounding tissue, seems to differ remarkably in the minds of our subjects, depending on the thought model that has been activated. Concretely, we see that prior activation of the balancing principle tends to influence a subjects' judgments towards consistency with the tumour-problem solution.

By the use of two different learning conditions (operationalised by providing „additive“ as well as „balancing“ learning examples of confluence) it could thus be shown that the judgment of ray confluence is not only schema-driven but depends largely on the activated mental contents „filling“ the representation of the problem, i.e. the thought model. This finding is valuable to our understanding about how people think about radiation confluence and to appreciate the mental demands involved in solving the classic tumour task. It also remained clear that there are substantial interindividual differences in judgments, which could indicate the existence of additional thought models.

2.5 Experiment 3

2.5.1 Aim

Experiment 2 tried to address the findings from Experiment 1, showing that a number of subjects did not view any difference in effectiveness or harm between one and four confluencing rays. The balancing thought model was identified to provide some explanation for this type thinking. We have proposed that the activation of this type of mental content, when thinking about confluence situations, is not only valid but in fact mandatory for the interpretation of the rays' harmfulness when solving the tumour problem according to Duncker's (1935) ideal solution. Using the same techniques as in Experiment 1 we elicited a shift in how our subjects judged the same confluence schema.

Returning to Figures 5 and 6 from Experiment 1, it is evident that in addition to those individuals who see the effects of a rays' confluence as strictly additive and those who judge it as balancing (or equilibrating), there are subjects who also describe one small ray to be less than four times less effective compared to four confluencing rays. When translated to the comparison between one big ray and four smaller rays, this means that one single big (four times stronger) ray would be less effective than four smaller (i.e. four times weaker) rays arranged in confluence. Naturally, we are reminded of the well-known Gestalt principle stating that the schematic whole is more than the simple sum of the single elements.

Hence, complementing the two confluence thought models of summation and balancing, we identified a third one, which shall be investigated in Experiment 3 (see Figure 3), using the same rational as in the previous experiments. Our hypothesis was analogous to the ones in the preceding experiments. Namely we expected that, the appropriate learning examples provided, we were able to create a shift in how the subjects judge the confluence effects applied to the tumour task. In accordance with the new thought model, one big ray needed to be viewed by more people as being less effective and less harmful compared to four confluencing small rays, as this was the case in the first two experiments. We intended to contrast the new results especially with those from Experiment 1. This is because the current „Gestaltist-like" principle states that the confluence effects are larger than the sum (i.e. additive thought model) of the single input.

2.5.2 Method

The latest principle of confluence was described to the subjects in the questionnaire's introduction as emerging from the properties of *distribution and angle of impact* (we shall refer to it simply as *distribution-based* thought model of confluence). The subjects were again provided with two learning examples, one being about the containment of a fire by use of the same amount of water per time period supplied from one side only compared to applying it from different sides simultaneously. They had to judge this type of confluence themselves in the learning task, rating the quality of sound experience of having one big loud speaker only, compared to having four 4-times less powerful loud speakers surrounding them (see Appendix 4.2 for the questionnaires).

With the exception of the new examples and the fact that the subjects in the current learning condition had to rate the effectiveness and harmfulness of *one single big* ray compared to four confluencing small rays, all other aspects of the questionnaire and procedure were the same as in Experiment 1 and 2.

2.5.3 Subjects

In all, 36 subjects filled in the questionnaire. All were recruited by use of an e-mail list and were enrolled as degree students at the University of Jyväskylä, Finland.

2.5.4 Results

As in Experiment 1 and 2, five subjects were excluded from the analysis due to severe violence of the basic task's logic in their responses (see explanation from Experiment 1). Data from 31 subjects were included in the analysis. The mean rating of the effectiveness ratio of one big ray compared to four small rays lay at m=1.18 with a standard deviation of s=.95. From this we can compute the subjects' judgment when comparing one small ray to four small rays, which is logically one fourth of the above ratio: m'=.29 and s'=.24. Just a reminder, this basic logic has been explicitly defined in the questionnaire and is inbuilt into the dispersion-convergence solution used by Gick and Holyoak (1980, 1983). In the following we will concentrate on the latter type of ratio in order to make comparisons between the experiments easier. For harmfulness the judgments were on average m'=.52, s'=.3.

Looking at the judgments on an individual level it is obvious that many subjects are of the opinion that four confluencing rays are more influential than one big ray. This is in line with our hypothesis stating that the effect of confluence is larger than the sum of the rays' single intensities. From Figure 10 we can conclude that 42 percent of the subjects judge a single small ray to be less than one fourth as effective as four confluencing rays. The relative amount of subjects with this opinion is thus clearly larger than in the previous confluence learning conditions (10 percent in Experiment 1 and nought percent in Experiment 2).

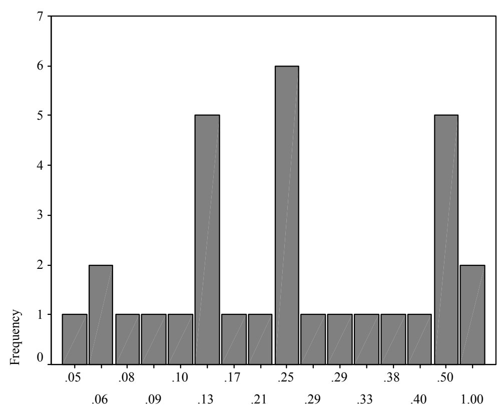


Figure 10: Effectiveness relationship: One big ray compared to 4 converging smaller rays

For harmfulness, the increase in judgments of the confluence situation as more effective than the summation of single rays is not apparent. Only 10 percent of the subjects in the current experiment are of this opinion (see Figure 11), which does not in any consistent way contrast the results from the previous experiments (19 percent in Experiment 1 and 3 percent in Experiment 2).

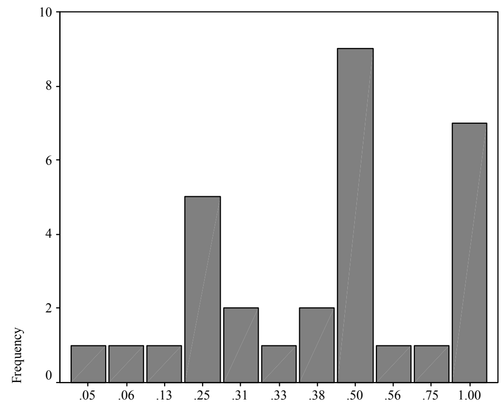


Figure 11: Harmfulness relationship: One big ray compared to 4 converging smaller rays

Interestingly a considerable amount of people are of the opinion that a single small ray is half as effective (N=5) and harmful (N=9) than four confluencing one. A comparable accumulation of this type of judgment was not found in the earlier experiments.

Testing of our main hypothesis in the form of t-tests of mean differences a tendency can be found according to which subjects view the effectiveness gain of confluence as bigger in the current learning condition compared to those subjects who studied additive confluence examples ($t(60)=-1.29$, $p_{(1\text{-tail})}=.1$). This tendency is in line with our hypothesis. For the judgment of the rays harmfulness the same hypothesis cannot however be substantiated. To the contrary it seems that the subjects in the current condition rate the increase in harmfulness as less dramatic than did the subjects in the additive condition ($t(60)=1.9$, $p_{(1\text{-tail})}<.05$). This inconsistency needed to be investigated further. For this we wanted to return to one of our two initial hypotheses, namely the one about the influence of the image of a ray (compact versus diverging). We remember that in Experiment 1, no substantiation was found for our assumption that subjects, who received rays displayed as diverging in nature, would see more harm done to the surrounding tissue when these rays are arranged in confluence. One explanation for this was that all subjects received examples for the additive thought model only, and that this restricted the full range of how confluence can be seen, taking all thought models into account. Therefore we wanted to test whether now, by including subjects from different confluence conditions, judgment differences regarding the harmfulness of radiation confluence could be found, and whether these would prove to be dependent on the type of ray representation displayed in the questionnaire.

In doing so, we first wanted to reassess the differences in harmfulness judgments given by our subjects in Experiment 1 and 3, separated for ray image conditions. An analysis of variance using the harmfulness judgments as a dependent variable and the two learning conditions from Experiment 1 and 3, as well as ray image conditions (compact and divergent) as factors was performed. The results showed that the interaction between the two factors is highly significant ($F(1)=7.75$, $p<.01$). As Figure 12 shows, only when the subjects that received pictorials with rays displayed in a compact form are compared, the harmfulness of confluence is seen as being smaller in Experiment 3 than in Experiment 1. Subjects from the diverging ray group in Experiment 1 and 3, on the other hand, behave according to our prediction.

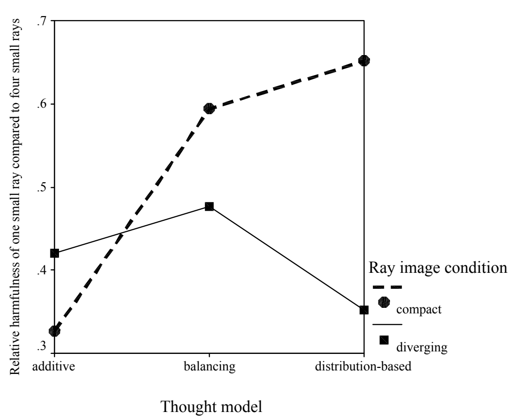


Figure 12: Interaction effects between ray image condition (compact vs. diverging) and thought models (additive (Experiment 1), balancing (Experiment 2), distribution-based (Experiment 3))

Along with the interaction effect, we now also seem to find evidence that the subjects in the diverging ray condition judge the increase in harmfulness (when using confluencing small rays instead of a single small ray) as much larger than do the subjects in the compact ray condition. To get a fuller picture, we included subjects from all three experiments (N=92) in the variance analysis, in order to look at whether significant mean differences in harmfulness judgments between ray image conditions could be found. With average judgments lying overall at m=.53 for the compact ray groups (N=48) and m=.42 for the diverging ray groups (n=44), we find a significant difference, not assuming homogeneity variance ($t(90)=1.72$, $p_{(1\text{-}tail)}<.05$; $F_{Levene}=1.76$, $p<.25$). This result suggests that the refutation of our hypothesis concerning ray image effects may be mainly due to the disparate judgments from the subjects in Experiment 1 (see Figure 12).

2.5.5 Discussion

In Experiment 3 a third thought model regarding confluence effects was investigated. Subjects in the current condition were not primed with examples displaying force confluence as equally effective, nor as the sum of the effectiveness of a single force, but instead as being more effective than the sum. This same thought model was applied to radiation by nearly half of the subjects, when judging the tumour destructive potential of confluence. Thus, the distribution-based thought model was used for effectiveness more than twice as often as the findings from the previous two conditions combined.

For the judgment of the confluencing rays' harmfulness to surrounding tissues, we did not find the results to be as clear. Only the subjects who received the rays displayed in a diverging manner answered according to the predictions based on the thought model. This made us return to the hypothesis about the ray image, left off after Experiment 1. A variance analysis checking for the effects of the pictorially suggested ray image on harmfulness judgments revealed significant differences between the ray image groups (compact vs. diverging), as well as an interaction effect between the thought model and the ray image condition. This interaction effect is most obvious in the counter-hypothetical answers provided by the subjects primed with the additive model. We currently have no explanation for this interaction effect.

3 General Discussion and conclusions

We focused on the current paper on two forms of mental content in a concrete context of Duncker's (1935) classic tumour problem: the representational image of a ray, and the thought model of confluence. We have shown that the judgment of confluence effects when using rays to destroy a tumour is not as straightforward as one might implicitly assume. The subjects differ significantly in their assessments concerning effectiveness and harmfulness of four confluencing small rays when being compared to a single ray situation. These differences are not self-evident when an explanation is based on elementary and schematic similarity, because the problems used in the experiments, did not differ in these respects.

Three thought models have been proposed to explain the judgment differences concerning one and the same confluence schema: the *additive* model, the *balancing* model, and the *distribution-based* model. Being considered the dominant thought model, the additive model was tested in Experiment 1, with the subsequent models being developed on the grounds of its results and tested in Experiment 2 and 3. The general idea was that the subjects would judge a rays confluence in consistency with the learned thought model, both with respect to the frequency of model-consistent answers provided, as well as the judgment average.

The results of the conducted three experiments provide valuable substantiation for the argument that the activation of different mental contents - although all equally meaningful in themselves - influences the construction representations of radiation confluence in disparate ways. Each experiment provided the subjects with distinct learning examples that stressed a certain mental content (or principle) in representing confluence. By use of pictorial illustrations we further suggested in each experiment two distinct contents when imagining rays. It was found that mental contents (i.e. of thought models and spatial images) activated in the representation of confluence and rays in the learning examples proved relevant for the representation of radiation confluence and the judgment of the solution to the tumour task. This means that the results of the experiments are a clear demonstration of content-based factors involved in causing interindividual and intercontextual judgment differences. With that there is also evidence for content-based transfer from the questionnaire's learning problems to the judgment task.

Further, the findings illustrate the difficulties that are involved when reasoners are required to mix different and conflicting thought models in the construction of a representation of a single task. Apperceiving the tumour problem with the nature of radiation confluence being additive with regards to its effectiveness in destroying a tumour, and non-additive with regards to the confluencing rays' harmfulness on the surrounding tissue may put subjects into a state of representational dissonance. It is obvious that if the researcher misses out on stating and including these content-related representational issues into his or her empirical design, deviations from theoretical assumptions may go undetected or remain inexplicable.

Our research also shows that the investigation into mental contents necessitates an emphasis in differential psychology over general psychology. It is empirically very common to think that statistical averages of collected data express essential issues in psychology. Indeed, this is mostly quite right. However, in content-based psychology the focus is often on individuals and groups, while overall averages tend to misrepresent them and lose information about diversity. Therefore, one is often better advised to focus on the variance and types of mental contents individuals and groups (e.g. intervention conditions) use in constructing their mental representations. This kind of information provides us with knowledge about what makes people apperceive the same physically perceivable stimuli differently.

The theoretical point of our research is clear. People have spatial images with certain mental contents and these contents explain the nature of transfer. Individuals differ from each other with regards to the thought models they use in apperceiving the presented tasks, and for this reason we get very different types of assessments. Some of them are appropriate and some are not, and transfer from the General to the tumour problem will respectively be positive or negative.

Our examples are spatial. They presuppose the organization of things in a mentally represented space. However, the crucial differences in representational contents lie in such concepts as divergence, confluence, addition, balancing and distribution-based representation. It is the contents of refered to by these concepts, which human apperception uses when constructing mental representations: Contents are the essence of representation (Saariluoma, 1997).

Thus, it is the analysis of contents that is essential in content-based explanations. One must be able to state why and how the transfer-related experience and reasoning of person A is different from person B. Or why person A shows a positive transfer in this situation but not in another. In order to do this, one can refer to different schemata or the use of different procedural elements in the transfer process. Here, we argue that eventually the real mental essence of these differences is best expressed in content-based theory language. A schema is by definition of the concept void of content and therefore it is not powerful enough to explain the transfer in the case where similarity is based on the contents of the schema. Rather, it is the *contents* of notions like confluence that explain the nature of transfer. It is by this highly valuable apperception-based approach that we seek to achieve a better idea about the psychological phenomena such as reasoning, transfer, or everyday experiences.

4 Appendix

4.1 The General-tumour story analogy

4.1.1 The General story (Gick and Holyoak, 1980, 1983)

A small country was ruled from a strong fortress by a ruthless dictator. The fortress was situated in the middle if the country, surrounded by farms and villages. Many roads led to the fortress through the countryside like spokes on a wheel. A rebel general vowed to capture the fortress and free the country of the dictator. The general knew that an attack by his entire army would capture the fortress. He gathered his army at the head of one of the roads, ready to launch a full-scale direct attack. However, the general then learned that the dictator had planted mines on each of the roads. The mines were set so that small bodies of men could pass over them safely, since the dictator needed to move his troops and workers to and from the fortress. However, any large force would detonate the mines. Not only would this blow up the road, but it would also destroy many neighboring villages. It therefore seemed impossible to capture the fortress.

However, the general devised a simple plan. He divided his army into small groups and dispatched each group to the head of a different road. When all was ready he gave a signal and each group marched down a different road. Each group continued down its road to the fortress so that the entire army arrived together at the fortress at the same time. In this way, the general captured the fortress and overthrew the dictator.

4.1.2 The tumor problem (Duncker, 1935; Gick & Holyoak, 1983)

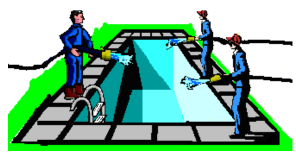
Suppose you are a doctor faced with a patient who has a malignant tumour in his/her stomach. It is impossible to operate on the patient, but unless the tumour is destroyed the patient will die. There is a kind of ray that can be used to destroy the tumour. If the rays reach the tumour all at ones at a sufficiently high intensity, the tumour will be destroyed. Unfortunately, at this intensity, the healthy tissue that the rays pass through on the way to the tumour will also be destroyed. At lower intensities the rays are harmless to the healthy tissue, but they will not affect the tumour either. What type of procedure might be used to destroy the tumour with the rays, and at the same time avoid destroying the healthy tissue.

4.2 The questionnaires (translations into English)

4.2.1 The learning section for the additive thought model

In many areas of our life we find the simple principle of *confluence*. „*Augmentation of simultaneously applied force*" is an important feature of confluence.

For example: When filling a swimming pool in our garden, we may use more than one hose at a time in order to increase the amount of water that can be supplied to the pool at a time (confluence). This means, of course, that the water level of the basin is after an equal time period higher than what it would be with only one hose (non-confluence).



As another example we may take humanitarian aid supplies delivered to devastated cities in Afghanistan. Especially in mountainous regions it was only possible to bring small loads of relief to the cities by using only one, usually mined, supply route (non-confluence). Emergency strategies were to have different humanitarian convoys' role into the needy cities from different directions simultaneously (confluence).

We all also have an understanding of what it means to lift a very heavy laundry machine alone, compared to having friends helping us.

What is according to you the lifting power of four comparably strong individuals, compared to one person alone? Choose your own anchor value for one person's lifting power (e.g. '1', '10', or any other you wish) and mark in relation to this the resulting value for the four people lifting together (any type of real number like '$^{1}/_{10}$', '0.8', '1.3', '$3^{1}/_{5}$', '10', ..., depending on your anchor value).

 set as ____	 ____

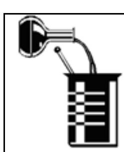
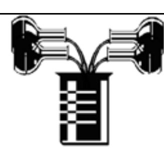
4.2.2 The learning section for the balancing thought model

In many areas of our life we find the simple principle of *confluence.* „*Threshold maintenance*“ is an important feature of confluence.

For example: Some modern bathtubs may (for aesthetical reasons) have more than one water tap. From each water trickles at 30º Celsius (confluence), keeping the whole tub at a constant temperature of 30º. In essence, the water is of course not necessarily colder in a conventional bathtub with only one tap (non-confluence).

Similarly, in a public sauna there may be two sauna ovens next to each other (confluenec). Yet obviously, the temperature does not rise to twice the maximal temperature one would reach with one oven (non-confluence), even when both are running at full power. Rather it stays roughly at the same temperature.

We all also have an understanding of mixing liquids. Liquid concentration (i.e. percentage of a critical substance in a liquid) is a very important factor in chemical laboratory work. What is according to you the resulting concentration of liquids from four containers mixed together, when each supply container contains a liquid with an equal concentration of a critical substance. Choose your own anchor concentration value for the liquid supplied by one container (e.g. '1' for 1%, '10' for 10%, or any other you wish) and then mark in relation to this the resulting value for the mixture concentration (any type of real number like '$^{1}/_{10}$', '0.8', '1.3', '3$^{1}/_{5}$', '10', ..., depending on your anchor value).

 set as ____	 ____

4.2.3 The learning section for the distribution-based thought model

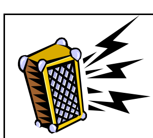
In many areas of our life we find the simple principle of *confluence*. „*Distribution and angle of impact"* is an important feature of confluence.

For example: When extinguishing a fire, the fire brigade may use many smaller hoses fed by one water tank instead of one big hose. By converging the water cannons onto the fire source (confluence), better control of the fire can be exerted than by using the same amount of water fed through one hose (non-confluence).

As another example we may compare the use of a hair-drying bell, where air is blown from different directions onto our head (confluence), to one more powerful hair dryer that is fixed and blows hot air from one direction only (non-confluence).

We all also have some experience with the arrangement of loud speakers in a room. Obviously it may make a qualitative difference whether we have one loud speaker blasting at us from one side, or 4 smaller loudspeakers, with overall equal output power, distributed in our room, so that they surround us.

What is according to you the sound experience when dividing the output power of one big loudspeaker into four smaller ones, arranged in a surround manner. Choose your own anchor value for on big loudspeaker (e.g. '1', '10', or any other you wish) and mark in relation to this the resulting value for the confluencing smaller loudspeakers (any type of real number like '$^{1}/_{10}$', '0.8', '1.3', '$3^{1}/_{5}$', '10', ..., depending on your anchor value).

 set as ____	 ____

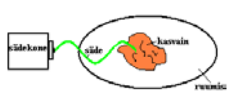
4.2.4 The transfer task

A final confluence example: In cancer medicine, researchers already realised far back that the potential hazardous effects of radiation could be used to destroy tumour cells in places where traditional surgery is not suitable.

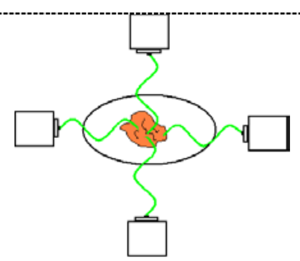
For the medical application of radiation to tumours the rays need to be strong enough to destroy tumour cells. The problem with this is that the cells of healthy tissue surrounding the cancerous area may then be just as negatively affected by it as the cells of the tumour.

Suppose you were a cancer researcher at the time, making judgments about the *effectiveness* (in destroying tumour cells) and *harmfulness* (on healthy cells surrounding the tumour) of two confluence techniques *A* and *B*.

A- The ray's intensity can be diminished to such a degree that it is largely harmless to healthy tissue (e.g. four times „smaller rays"). Then the tumour cells are also not affected. The idea of technique A is to use four smaller ray beams simultaneously, shot from different angles into the direction of the tumour. Judge below the *effectiveness* of *four smaller* rays in relation to *one small* ray. Judge also the potential *harm* of the confluence technique to the surrounding cells, again in relation to one small ray.



Effectiveness: set as ___ Harmfulness: set as ___



Effectiveness: ____ Harmfulness: ____

(Notice: In the questionnaire of the distribution-based condition, instead of the single small ray option, the confluence was judged in relation to a single big ray (four times thicker line was displayed))

5 References

Anderson, J. R. (1983). *The Architecture of Cognition.* Cambridge, MA: Harvard University Press.

Anderson, J. R. (1985). *Cognitive Psychology and its implications.* New York: W. H. Freeman & Co.

Anderson, J. R. (1993). *Rules of the Mind.* Hillsdale, NJ: Lawrence Erlbaum Ass., Inc.

Campione, J. C., Brown, A. L., & Ferrara, R. A. (1982). Mental retardation and intelligence. In R. J. Sternberg (Ed.), *Handbook on human intelligence* (pp. 392-490). New York: Cambridge University Press.

Detterman, D. K. (1993). The case for prosecution: Transfer as an epiphenomenon. In D. K. Detterman, & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 39-67). Stamford, CT: Ablex Publishing Corp.

Duncker, K. (1935). *Zur Psychologie des produktiven Denkens.* Berlin: Springer.

Gick, M., & Holyoak, K. (1980). Analogical problem solving. *Cognitive Psychology, 12*, 306-355.

Gick, M., & Holyoak, K. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1-38.

Issing, L. J., Hannemann, J., & Haack, J. (1989). Visualization by pictorial analogies in understanding expository text. In H. Mandl, & J. R. Levin (Eds.), *Knowledge acquisition from text and picture* (pp. 195-214). Amsterdam: Elsevier.

Kant, I. (1781). *Kritik der reinen Vernunft.* Stuttgart: Philip Reclam.

Keane, M. (1987). On retrieving analogues when solving problems. *The Quarterly Jounral of Experimental Psychology, 39*A, 29-41.

Maarttola, I., & Saariluoma, P. (2002). Error risks and contradictory decision desires in urban planning. *Design Studies, 23*(5), 455-472.

Moran, T. P. (1983). Getting into a system: external-internal task mapping analysis. In *Proceedings CHI'83 Human Factors in Computing Systems.* Boston.

Novick, L. R. (1990). Representational tansfer in problem solving. *Psychological Science, 1*(2), 128-132.

Polson, P. G., & Kieras, D. E. (1985). *A quantitative model of the learning and performance of text editing knowledge*. Proceedings of the CHI '85 conference on Human factors in computing systems.

Reed, S. K. (1987). A structure-mapping model for word problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 124-139.

Saariluoma, P. (1990). Apperception and restructuring in chess players' problem solving. In K. J. Gilhooly, M. T. G. Keane, R. Logie, & G. Erdos (Eds.), *Lines of thinking: Reflections on the psychology of thought, Vol. 2: Skills, emotion, creative processes, individual differences and teaching thinking* (pp. 41-57). Oxford, Eng.: John Wiley & Sons.

Saariluoma, P. (1992). Error in chess: The apperception-restructuring view *Psychological Research/Psychologische Forschung, 54*(1), 17-26.

Saariluoma, P. (1995). *Chess players' thinking.* London: Routledge.

Saariluoma, P. (1997). *Foundational analysis*. London: Routledge.

Saariluoma, P. (2001). Chess and content-orientated psychology of thinking. *Psicològica, 22*, 143-164.

Saariluoma, P. (2002). *Thinking in work life.* (In Finnish). Porvoo: WSOY.

Saariluoma, P. (2003). Apperception, content-based psychology and design. In U. Lindemann (Ed.), *Human behavious in design* (pp. 72-78). Berlin: Springer.

Saariluoma, P., & Hohlfeld, I. (1994). Chess players' long range planning. *European Journal of Cognitive Psychology, 6*, 1-12.

Saariluoma, P., & Kalakoski, V. (1997). Skilled imagery and long-term working memory. *American Journal of Psychology, 110*, 177-201.

Saariluoma, P., & Kalakoski, V. (1998). Apperception and imagery in blindfold chess. *Memory, 6*, 67-90.

Saariluoma, P., & Maarttola, I. (In press). Stumbling blocks in novice building design. *Journal of Architectural and Planning Research*.

Simon, H. A., & Reed, S. K. (1976). Modelling strategy shifts in a problem solving task. *Cognitive Psychology, 8*, 86-97.

Singley, M. K., & Anderson, J. R. (1985). The transfer of text-editing skill. *International Journal of Man-Machine Studies, 22*, 403-423.

Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill.* Cambridge, Ma: Harvard University Press.

Stevens, S. S. (1961). Procedure for Calculating Loudness: Mark VI. *Journal of the Acoustical Society of America, 33*(11), 1577-1585.

Stout, G. (1896). *Analytic psychology.* New York: Macmillan.

Thorndike, E. L., & Woodworth, R. S. (1901a). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review, 8*, 247-261.

Thorndike, E. L., & Woodworth, R. S. (1901b). The influence of improvement in one mental function upon the efficiency of other functions: The estimation of magnitudes. *Psychological Review, 8*, 384-395.

Thorndike, E. L., & Woodworth, R. S. (1901c). The influence of improvement in one mental function upon the efficiency of other functions: Functions involving attention, observation and discrimintation. *Psychological Review, 8*, 553-564.

Wundt, W. (1880). *Logik I.* Stuttgart: Ferdinand Enke.