Running head: MENTAL CONTENTS IN TRANSFER

# Mental Contents in Transfer

Sacha Helfenstein and Pertti Saariluoma[*]

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

[*] corresponding author

Abstract

The purpose of this paper is to develop theoretical concepts for mental content-based investigations and explanations in the psychology of human thinking in general and transfer in specific. The schema-based analysis of transfer postulates that the solving of an earlier problem can influence a person's future behaviour only to the degree of schematic similarity between the primary and the secondary problem solution. The reported content-based investigation, using Duncker's (1935) classic tumour task, shows however, that the *contents* of schemata cause an essential variation in a person's mental representation and judgement of radiation confluence effects. The study identified and assessed in separate experiments the influence of 3 different thought models (*additive*, *balancing*, and *distribution-based*) and 2 distinct types of spatial images of rays (*compact* vs. *diverging*). Whereas differences in the central tendencies of the judgements could expose some of the priming effects between experimental conditions, the core of the content-based analysis was based on the identification of distinct groups of participants displaying a certain type of judgement. These differences reflected the contrasts between the thought models and the ray image content. They substantiated the claim that a schema-based analysis of transfer and reasoning in general, alone is not sufficient enough to explain inter-individual and inter-contextual differences that are based on distinctive mental content in reasoners' apperception.

# Introduction

It is natural to assume that information contents in mental representations can explain human behaviour. Research approaches, which apply the concepts of mental contents as the explanatory framework in investigating psychological phenomena can be called content-based (Saariluoma, 2002, 2003). There are a number of research topics such as economic and design thinking, in which it has been proved proper to employ a content-based approach (Saariluoma, 1990, 1992, 1995, 1997, 2001, 2002, 2003; Saariluoma & Hohlfeld, 1994; Saariluoma & Kalakoski, 1997, 1998; Saariluoma & Maarttola, 2003). Consequently, it is warranted to continue to systematically develop the basic methods and concepts. In doing so, the adaptation of traditional experimental paradigms of thinking has been of great value.

A key issue in content-based psychological research is to understand what kinds of psychological processes are responsible for the information contents of our mental representations. An intuitive but clearly insufficient answer is that the origin of mental content lies in perceptual experience alone. This empiricist explanation fails to capture a number of non-perceivable kinds, which are so evidently incorporated within our representations. Humans can represent such content as infinity, future, past, possible, allowed, constitution, n-dimensional spaces, all of which cannot be perceived. It is also very commonplace that two people “see” identical objects very differently. Expert chess players or architects, for example, represent the same perceptual stimuli in a very different fashion than do novices. For these reasons, it is essential to make a distinction between perception and *apperception* (Saariluoma, 1990, 1995, 2003; Saariluoma & Maarttola, 2003). Historically the concept of apperception has been used in different ways, but we define it as the process of constructing mental representations (Saariluoma 1990; see, for history of the concept, Brentano, 1874; Kant, 1781/1956; Leibniz, 1704/1989; Stout, 1896; Wundt, 1896/1920). Our own empirical research has demonstrated that content-based thinking is essential in analysing what kind of principles apperception follows in selecting and integrating information contents into consistent mental representations (Saariluoma 1990, 1992, 1995; Saariluoma & Maarttola, 2003).

It is our ascertainment that the apperception- and content-based theoretical concepts would have sufficient power of expression (Saariluoma, 1997) to be used in the investigation of all aspects of the construction of information contents in mental representations. In this paper we empirically demonstrate our point by contrasting the concepts of *schema* and *thought model*. We study our assumption that content-based concepts would improve the power of conceptual analysis in transfer experiments compared to schema-based thinking.

Schema is a traditional notion that is widely used to express mental content and it is a well-known argumentation that people selectively encode the world by means of schemata (Neisser, 1976). Transfer refers to the case where the mastery of one situation makes the subsequent handling of another easier or more difficult (Ellis, 1965; Saariluoma, 1992; Singley & Anderson, 1985, 1989; Woodworth, 1951). In schema-based transfer experiments it is assumed that by confronting a person to a primary problem's solution, a schema is tought, which can be applied when solving a secondary or target problem (Gick & Holyoak 1980, 1983; Reed, 1993). The problem with the schema notion, however, is that it does not really capture the essence of mental content; in fact, it is by definition void of content. If it is possible to illustrate that the *contents* of schemata cause an essential variation in behavior and, respectively, mental representations of individuals, it is reasonable to conclude that these differences should be explained in terms of mental contents.

In our experiments, we want to spotlight our participants' representation of rays and their confluence. Duncker (1935), as well as Gick and Holyoak (1980, 1983) tacitly assumed that their participants imagined medical rays in a compact, laser-like fashion and that they understood the effectiveness of the rays' confluence for destroying the tumour in an additive way (see Figure 1). Additive confluence means that the energy of four smaller rays converged on a single point equals the energy of a single, four times more powerful ray. In contrast to the conceptualisation of effectiveness, Duncker ascertained that the harmfulness of the confluence technique to the surrounding tissue would remain unchanged (non-additive), when compared to that of a small single ray. It is obvious that the contradictory potential of these experimental presuppositions needed to be investigated.

----------

Insert Figure 1 somewhere here

----------

We devised two preparatory studies and built our experimental paradigm on their empirical results and our content-based theoretical considerations. In the first study we wanted to get more insight into why the majority of participants fail to implement the schema analogy envisioned by Gick and Holyoak (1980, 1983) when solving the tumour problem. The second one served us to obtain a better understanding about a person's implicit representation of a ray.

The core of our investigation comprised of three consecutive questionnaire-based experiments. With regard to the conceptualisation of the confluence schema, the purpose of the first was to substantiate the value of the additive thought model when judging the ray effects and to identify disparate types of judgment. In the two following experiments we then wanted to characterize separately and assess alternative thought models, which underlie the judgment variations found in Experiment 1. In all three experiments the same priming-transfer design was used, with only the learning examples used for priming being altered according to the content of the currently studied thought model of confluence. During transfer, all participants had to judge the effects of the rays' confluence for the same tumour task.

While the transfer part of the experiment was not manipulated between experiments, two distinct ray illustration versions were used within each of the experiments. The intention of this was to look into the potential representational predicament concerning Duncker's (1935) presupposition that people imagine rays in a compact, laser-like fashion.

Generally, we intended to demonstrate that the content-based factors of the confluence thought model and the ray image governs the apperception process of constructing a mental representation of the radiation situation as a whole. We expected that our manipulations of the learning condition and the ray image would affect a person's reasoning and problem solving, and consequently, the nature of transfer.

## Experimentation

### *Preparatory Study 1*

#### *Stimuli and procedure.*

The participants read one of two devised critical learning stories (one of which was the original General problem with solution used by Gick & Holyoak, 1980, 1983) in between two others, here irrelevant stories, in the context of a memory recall experiment. Immediately after its completion, the participants were asked to participate in a separate experiment and guided to a different room. Here they were confronted with Gick and Holyoak's version of the tumour problem and urged to talk-aloud and provide a deliberate amount of solutions during a five minute time period. For those who did not spontaneously implement the confluence solution the experimenter increased the salience of the analogy between the critical learning story in the first experiment and the tumour problem by presenting a paper with a preprinted image of a fortress and converging roads leading towards it next to a sheet with a tumour patient.

*Participants.*

We randomly recruited 8 participants (6 female, 2 male students; age range: 22-27), through mailing lists at the University of Jyväskylä, Finland, and rewarded them with a movie ticket.

*Results and discussion.*

With only one student coming up with the critical confluence solution prior to the hint, Gick and Holyoak's (1980, 1983) findings about poor spontaneous transfer were replicated. However, the analysis of the participants' solution rich talk-aloud reports indicated to us that this may be less a problem of the lack of transfer or spontaneity as such, and more of an inappropriateness of the presumptions made about the mental relationship between the learning and the transfer situation. Moreover, it appeared that the main obstacles for General- to-tumour story transfer might be largely inbuilt into the tumour problem itself. The participants argued to know too little about the properties of radiation, in order to reason by analogy. When they commented on their intuitive and naive representations of rays and their confluence (e.g. “chaotic”, “diffuse”, “diverging”, “hard to control”, “non-additive”) it became evident that many of these mentioned points imposed a real obstacle in problem solving and may even be irreconcilable with the type of schematic transfer assumed in the Gick and Holyoak (1980, 1983) task.

*Preparatory Study 2*

*Stimuli and procedure.*

We administered a questionnaire where the participants were required to explicitly point out their imagination of (medical) radiation by choosing from one of three given representational options (laser-like, diverging, diffuse) or provide their own.

*Participants.*

The study was conducted in the context of a lecture at the Chydenius Institute in Kokkola, Finland, with 15 students (4 female, 11 male students; age range: 20-50).

*Results and discussion.*

Interestingly we found that while three quarters of the students reported to imagine rays in a radial diverging fashion, only *one* person selected the image of a compact ray beam. This result reaffirmed our suggestion that Duncker's (1935) ideal laser-like type of ray-representation may not reflect a persons' naive conception, and that this representational factor may critically influence the way a participant judges the effects of the ray confluence technique.

*Experiment 1*

As was pointed out in the introduction, the additive thought model plays a crucial part when judging the rays' effectiveness when destroying the tumour. The preparatory, qualitative analysis of our participants' solutions to the tumour problem has also shown that there was great ambiguity about how to conceptualise the confluence effects. Our first experiment was therefore designed to prime the participants with the additive model of confluence and to assess the consistency in which they adopted this way of thinking when judging the radiation problem.

In both our preparatory studies we have also found considerable uncertainty concerning the ray image representation. Above all we discovered that peoples' intuitive representations differ sharply from the presumptions made by Duncker (1935) and Gick and Holyoak (1980, 1983), and that this may jeopardise the conception of the confluence solution. We decided to compare the compact image of ray with the diverging representation, because the latter was the most frequent choice in Preparatory Study 2. Together with the talk-aloud reports from Preparatory Study 1 we built the hypothesis that a person will rate the harm-diminishing effect of radiation confluence (instead of using one big ray, for instance) as less pronounced when the ray-beam is suggested as diverging in nature, compared to a compact image of a ray.

*Stimuli and procedure.*

A questionnaire containing two learning examples, a learning judgement task, and a transfer judgement problem was presented to the students. The intention of the introduction and learning section was to make the students familiar with the additive model of confluence. It was explained that one important effect of confluence is to augment forces in an additive fashion. Two learning examples and a learning judgement task demonstrated this idea using elements from outside the domain of radiation. The first example used the idea of water confluence to fill up a swimming pool, the second drew on the principle of multiplying the amount of simultaneously transported humanitarian aid supplies to a crisis area by employing more than one convoy starting from different storage centres around the area.

In the learning judgement task the students needed to provide their own estimation of the confluence effects for a rather simple control task in order to familiarise them with the relative magnitude estimation technique, also used in the transfer task. This method is described in detail by Stevens (1961) to measure stimuli perception. The matter of judgement was an evaluation of the lifting power of a single person compared to four equally strong people combined.

The transfer task followed in essence as a fourth example of confluence, however, the participants were not informed that the confluence of rays obey the same additive law as their learning examples did. Here the participants were urged to give a relative estimation of the effectiveness and harmfulness of confluencing rays for destroying a tumour and preserving the surrounding healthy tissue likewise. As in the learning judgement task, the participants 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 for the confluence situation (see rows in Figure 2). To test a possible influence of ray image representation on confluence judgement, two experimental conditions were created by distributing slightly different versions of the questionnaire among the students. One group received a questionnaire that had the rays displayed in a diverging manner, the other with compact ray images (see columns in Figure 2). The distinction between the two was entirely of a latently suggestive nature, i.e. nothing in the text explicitly underlined the image condition.

An alternative option would have been to let the participants provide their own intuitive representations of a ray. Aside from the operational challenge of such an approach and the need to lengthen the questionnaire, it already became clear from the priorly conducted qualitative interviews that this conscious reflection upon ones own implicit representation poses great difficulty on most participants.

----------

Insert Figure 2 somewhere here

----------

*Participants.*

A total of 36 participants, all enrolled as degree students at the University of Jyväskylä, Finland, filled in the questionnaire. The students were randomly recruited from a lecture course (course paticipants' age range: 19-45, $M$ = 24, $SD$ = 4; sex: 25.5% female; 74.5% male) during the autumn term of 2003.

We excluded 5 participants, whose answers seemed to violate some basic logic (i.e. one small ray was more harmful than four). The reason for their rejection was that we could not decide whether their answers represent a genuine, but very deviant (or irrational) interpretation of the task, whether these participants simply misrepresented the task illustration, or whether they generally behaved in a non-cooperative way. The fact that 4 of these participants also deviated strongly from all others when judging the learning task (e.g. the lifting power of four people was noted as 1/3 compared to one person) indicated to us that they may also have misunderstood the assessment technique.

*Results.*

The frequency charts (Figure 3) depicting the judgements of the 31 participants show that, with the medians laying at .25, the majority of the students (20 for effectiveness, 16 for harmfulness) judged one small ray to be four times less effective and harmful than four of the same small rays (additive model). Four persons viewed the single ray situation as equally effective as the confluence situation (relation = 1), 4 gave an intermediate judgement where four confluencing rays increase the effectiveness somewhat but not fully proportional (relation between .25 and 1), and 3 considered the effectiveness of four rays as more pronounced than the additive model suggests (relation < .25). This variability in the judgments produced in average a significant deviation from the ideal additive model with a value of .25, $t(30) = 2.61$, $p < .05$.

----------

Insert Figure 3 somewhere here

----------

As expressed above, a majority also judged the harmfulness relationship according to the additive thought model. Four participants considered confluencing rays as equally harmful as one ray alone (Duncker-solution), 5 gave intermediate assessments, and 6 estimated the increase in harmfulness as greater than the summation of single harm effects.

Furthermore we wanted to explore the harmfulness ratios as dependent on the received ray image illustration. This contrast did not produce the expected result. The participants in the diverging ray condition ($n = 16$, $M = .42$, $SD = .32$) judged the increase in harm, when using four rays in a confluencing manner, on average not significantly different from the participants in the compact ray condition ($n$ = 15, $M$ = .33, $SD$ = .24), $t(29) = -.93$, $p_{(1\text{-tail})} > .1$. However, the participants' judgement in the diverging ray condition were also somewhat more heterogeneous ($F_{\text{Levene}} = 2.39$, $p = .133$). From the analysis of the interindividual distribution it became apparent that, contrary to our hypothesis, only in the diverging ray image condition did some students tend towards estimating the harmfulness relation as 1.

*Discussion.*

The results of Experiment 1 show that the additive thought model was the dominant model when judging the effects of applying a confluence schema. We found that on average the students judged four confluencing small rays as more effective but also more harmful than one single small ray. The latter challenges the ideal understanding presupposed in the solution to Duncker's (1935) tumour task and it is suggested that one reason for this result 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 the participants in Experiment 1 would be able to appreciate the functional meaningfulness of the confluence solution for diminishing harm.

The different pictorial presentation of rays used in the questionnaire (diverging vs. compact) did not, however, produce the expected effects on the harmfulness judgement. It may be that the additive confluence examples produced an over-homogenization of how people think about radiation confluence, or that our research assumption has been inappropriate. People may well interpret diverging rays as losing energy on the way to the tumour, which does not harm the surrounding tissue, but simply disperses into "space". It could also be possible that the latent suggestion provided in our illustrations were not distinctive, effective or reliable enough.

Last but not least we found that in spite of priming the participants with the additive thought model, a considerable number of them do not apply this content to the judgement of ray confluence. Obviously we need additional thought models to explain these interindividual differences, which were even more pronounced in the diverging compared to the compact ray image condition. One alternative thought model - already inbuilt into the Duncker solution - suggests that the effect of confluence does not exceed a single ray's effectiveness and harmfulness. A second alternative one must be responsible for such judgements where confluence effects exceed even the summation of single effects. Finally, we found such relation judgments that lay in between the results of applying the additive model and those where confluence is seen as irrelevant for the effects of radiation (we called them intermediate judgments; see also Figure 3).

*Experiment 2*

In the second experiment we wanted to make an investigative step towards a better understanding of the discovered variety in the effectiveness and harmfulness judgements. We needed to establish whether different types of judgements could be represented by distinct thought models. In contrast to Experiment 1, where we worked with additive learning examples of confluence, we now tried to find such learning examples where the effect of arranging input sources to create confluence are equal to the effect of a single input source.

We expected that, given appropriate learning examples, in Experiment 2 a greater proportion of people will judge the confluence effects in this manner, and that the average judgement ratios obtained will be significantly larger than the same ratios judged by participants primed with the additive confluence examples.

*Stimuli and procedure.*

For comparability reasons we chose the same procedures as in Experiment 1. We changed the introduction and learning section of the questionnaire where needed to emphasise the new thought model, but we decided to continue working with the same ray pictorials to create ray image representation conditions. 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. We elucidated this idea using two temperature confluence situations as learning examples. One stated that having only one compared to two or more taps supplying water with a temperature of 30 degrees Celsius to a bath tub does not change the resulting warmth of the bathing water; provided, of course, that the tub is well isolated. The same idea was also exemplified using the case of larger saunas with more than one oven.

A liquid concentration task was used in the learning judgement task, which the students had to judge for themselves. 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 are filled into a bowl the proportional content of sugar remains the same.

The continuation of the use of a simple learning judgement task to introduce the relative estimation technique was an important decision to make, especially because the data collected in Experiment 1 allowed for the suggestion that a substantial amount of people had problems with using this method. We decided, however, against the use of a standardised training task, such as the judgement of the relative lengths of lines (see Allard, 2001; Stevens, 1963), in order not to risk to bind the students' interpretation of the contrast between the two ray arrangements (single ray vs. confluence) to a certain type of thinking, which would interfere with our empirical intentions.

Finally, it is worth noting that the model description used in the questionnaire was a simplified one. 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 (see Figure 4).

----------

Insert Figure 4 somewhere here

----------

Accordingly, the concept of balance is used hereafter to describe the essence of the current thought model. However, since all our learning examples as well as the ray judgement task dealt with the confluence of equal input forces, it proved sound enough to use the simple model description in the questionnaire.

*Participants.*

Another 35 students enrolled in the same lecture course as those in Experiment 1 filled in the questionnaire. Like in Experiment 1, 5 participants needed to be excluded from the analysis due to suspicion of being non-cooperative or having misunderstood the logic of the assessment.

*Results.*

Looking at the central tendencies of the judgements provided by the 30 participants who have correctly completed the questionnaire, we found slight increases on the average ratios given compared to the results in Experiment 1 (effectiveness: $M$ = .42, $SD$ = .28; harmfulness: $M$ = .54, $SD$ = .36). This mean deviation from Experiment 1 proved significant only for harmfulness, $t(59) = -1.97$, $p_{(1\text{-tail})} < .05$, while homogeneity of variances cannot be assumed ($F_{\text{Levene}} = 9.855$, $p < .01$). It clearly appeared that our participants´ opinions about confluence effectiveness deviated not even more strongly from the additive judgements, $t(29) = 3.26$, $p < .01$, which is in line with our research assumption. At the same time the observed mean increase in harmfulness judgements pointed in the right direction, but remained clearly below the ratio of 1, ideally suggested by the balancing (or threshold maintenance) principle.

Looking at the frequencies charts (Figure 5) it is apparent that judgements below .25 (i.e. people that rate the effect of one small ray as less than one fourth of four confluencing small rays) have completely disappeared for effectiveness, and only one person makes such a judgement when considering the confluence's harmfulness.

----------

Insert Figure 5 somewhere here

----------

Further we found that for harmfulness there is a noticeable increase in answers that are in line with the balancing thought model (11 participants). This was not found for the judgement of effectiveness (5 participants), where additive judgements remained clearly the dominant way in thinking about the confluence effects (19 participants, compared to 14 for the harmfulness judgement). Similarly to Experiment 1, 4 to 6 people provided intermediate judgements.

Finally, we again failed to find a significant mean difference between the ray image condition groups for the judgement of the ray confluence's harmfulness, $t(28) = .88$, $p_{(\text{t-tail.})} = .19$. Also no difference in judgement variance was found between the groups ($F_{\text{Levene}} = 1.7$, $p = .2$).

*Discussion.*

Experiment 2 tried to address the findings from Experiment 1, showing that a number of participants did not view any difference in the effectiveness or harm between using one or four confluencing rays. The balancing thought model was identified as underlying this kind of thinking. Earlier we established 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 procedure as in Experiment 1, we elicited some shifts into how our participants judged confluence effects. The same confluence schema, judged in respect to the rays' harmfulness to the surrounding tissue, differed substantially in the minds of our participants, depending on the thought model that has been activated. Concretely, we found that prior activation of the balancing principle tends to influence a person's judgement towards consistency with the Duncker-solution. This was true for average measurements as well as for the individual distribution. The results from our second experiment, however, also demonstrate the robustness of the additive thought model, especially in judging the ray confluence's effectiveness with respect to destroying a tumour. Here, only a small and nonsignificant mean difference could be found between prior activation of the mental content of the additive model or the one of balance.

An interesting difference to Experiment 1 was also that the use of the third thought model, the one judging the effects of confluence as greater as the sum of the single rays, had nearly vanished. Experiment 3 shall try to address the essence and effects of that thought model.

*Experiment 3*

It became evident from Figure 3 in Experiment 1, that in addition to those individuals who see the effects of a rays' confluence as strictly additive, those who judge it as balancing, and those who give intermediate judgements, there are people who consider one small ray to be less than four times less effective compared to four confluencing rays. Our next challenge lay therefore in the conceptualisation and operationalisation of the latter thought model.

We expected that in accordance with the new thought model and the learning examples provided (see below), confluence effects needed to be viewed by more people as exceeding the summation of single effects. Figure 6 illustrates the idea of the current thought model in an analogy to Figures 1 and 4.

----------

Insert Figure 6 somewhere here

----------

Due to their communality of being linked to the concept of summation, we especially intended to contrast the new results with those from Experiment 1 (i.e. additive thought model).

*Stimuli and procedure.*

Based on the idea of the learning examples the latest principle of confluence was described to the participants in the questionnaire's introduction as emerging from the properties of *distribution and the angle of impact* (we shall refer to it simply as *distribution-based* thought model of confluence).

One of the learning examples dealt with 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. The second learning example used the same logic for the example of drying hair with one big or four smaller hairdryers. In the learning judgement task the participants had to estimate the confluence effects themselves by 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.

Because we employed the relative estimation technique, the transfer task could be readily transformed from the previously used small ray-relations into a proportional comparison between *one single big ray* and four confluencing small rays. Restated in this fashion, the thought model read that one single big (four times stronger) ray would be less effective than four smaller (i.e. four times weaker) rays arranged in confluence. This is naturally very similar to the well-known Gestalt principle stating that the schematic whole is more than the simple sum of the single elements.

In all other aspects of the questionnaire and procedure, other than those mentioned, we used the same methods as in Experiment 1 and 2.

*Participants.*

We again asked different students enrolled in the same university lecture course as those in Experiment 1 and 2 to fill in the questionnaire. A total of 35 participants participated in Experiment 3, of which 4 were excluded from the analysis due to our suspicion about non-cooperative behaviour and their illogical deviations from the others in judging the learning task (see also explanation from Experiment 1).

*Results.*

We have recoded the ratio judgements of Experiment 3 in order to be able to display the results in a way that is consistent with the logic of Experiment 1 and 2. I.e. the ratios depict the effectivity of a single small ray in proportion to four confluencing small rays[2]. The interindividual distribution of the effectiveness judgements (Figure 7) shows that 42 percent of the participants are of the opinion that four confluencing rays are more influential than just the sum of the rays' powers. The relative amount of students with this opinion is thus clearly larger than in the previous confluence learning conditions (Ten percent in Experiment 1, and zero percent in Experiment 2).

----------

Insert Figure 7 somewhere here

----------

For harmfulness, an increase in distribution-based judgements of confluence is not as apparent. Ten percent of the participants in the current experiment are of this opinion (see Figure 7), which contrasts the results of the previous experiment (3 percent), but does not in a model-consistent way differ from the findings in Experiment 1 (19 percent).

Interestingly a considerable amount of people are of the opinion that a single small ray is half as effective (5 participants) and harmful (9) than four confluencing ones. A comparable accumulation of this type of intermediate judgement was not found in the earlier experiments. Balancing-based judgements were given by two and seven people for effectiveness and harmfulness respectively.

The average rating of the effectiveness in Experiment 3 was $M = .29$. For harmfulness the judgement was on average $M = .52$. T-tests of mean differences of these judgements to the results in Experiment 1 show a tendency according to which the current participants view the effectiveness gains of confluence as larger compared to those who studied the additive confluence examples, $t(60) = -1.29$, $p_{(1\text{-tail})} = .1$. This is in line with our hypothesis. For the judgement of the rays' harmfulness the same hypothesis cannot be substantiated. To the contrary it seems that the students in the current condition rate the increase in harmfulness as less dramatic than did those in the additive condition, $t(60) = 1.9$, $p_{(1\text{-tail})} < .05$. This is an interesting inconsistency, which needed additional analysis (see below).

In contrast to the findings of the previous experiments, we further discovered in Experiment 3 a highly significant and hypothesis-consistent mean difference in harmfulness judgements between the ray image conditions (compact vs. diverging) groups, $t(29) = 3.16$, $p < .01$ ($F_{\text{Levene}} = 3.54$, $p = .07$).

An interesting question arose from the previous two results: "Is there a possible interaction effect to be found between a thought model and a ray image when judging ray confluence's harmfulness?" By including all participants ($n = 92$) primed with different confluence models we first wanted to test whether judgement differences regarding the harmfulness of radiation confluence could be found, which depend on the type of ray representation displayed in the questionnaire. An analysis of variance was performed, using the harmfulness judgements as dependent variable and the different thought models as well as the ray image conditions (compact and divergent) as factors.

----------

Insert Figure 8 somewhere here

----------

The results showed that the interaction between the two factors is significant, $F(2,86) = 3.16$, $p < .05$, and suggests that the earlier refutation of our hypothesis concerning ray image effects may be mainly due to the disparate judgements from the participants in Experiment 1. Figure 8 also shows that, only when the participants who received pictorials with rays displayed in a compact form are compared, the confluence's increase in harmfulness is seen as being smallest in Experiment 3. Participants from the diverging ray group, on the other hand, behave according to our prediction.

Along with the interaction effect, we also found evidence for the tendency that the participants in the diverging ray condition judge the increase in harmfulness (when using confluencing small rays instead of a single small ray) on average as larger than the participants in the compact ray condition, $F(1,86) = 2.8$, $p < .1$.

*Discussion.*

In Experiment 3 the effects of a third thought model regarding the confluence effects was investigated. In contrast to the earlier examples where the impact of confluence was displayed in a summative or balancing way, participants in the current condition were primed with a model portraying the confluence's effect as exceeding the sum of single forces. This thought model was applied to the radiation problem by nearly half of the participants, when judging the tumour destructive potential of confluence. Thus, the distribution-based thought model of effectiveness was used more than twice as often compared to findings from the previous two conditions combined.

For the judgement of the confluencing rays' harmfulness to surrounding tissues, the results were less consistent. Only the students who received the rays displayed in a diverging manner answered in line with the new thought model. This interaction effect between the ray image condition and the received thought model proved substantial when including participants from all three experiments in a variance analysis. It also reflected the almost counter-hypothetical answers provided by the students receiving different ray image pictorials in the previous experiments. Currently we have no explanation for this interaction effect, but it reveals the relevance of the process of integrating mental contents in apperception.

Just as in the first two experiments, we also found in Experiment 3, that, although a clear shift in content can be evoked, students use the full variety of models when judging the effects of confluence in the radiation task. Students primed with the distribution-based model appeared to provide exceptionally often intermediate judgements (ratings in between additive and balancing ones). In particular the judgement considering the confluence of four rays as being twice as effective compared to a single was remarkably frequent. It is therefore obvious that the thought models identified and studied in the three experiments do not exhaustively represent the full variety of content involved in the thinking of the confluence schema. There may be one or even several additional thought models explaining intermediate and other kinds of judgements, but these lie within the variability range of the models studied so far and do not offer any supplementary value for the conclusions we are about to make regarding our research questions. Hence, no further experiments need to be conducted at this stage.

## General Discussion

In our experiments, we varied systematically two dimensions of the tumour radiation task used by Duncker (1935) and Gick and Holyoak (1980, 1983). Firstly, we primed the participants with three thought models, different in content, related to the confluence schema. They were (a) the *additive,* assuming a summation of a series of single inputs, (b) the *balancing,* assuming an average over single inputs, and (c) the *distribution-based,* suggesting that confluence exceeds the sum of the single inputs. Secondly, we manipulated the ray image conditions, i.e., whether the ray is diverging or converging.

The general idea of our study was that schematic reasoning is influenced by the contents of our thoughts. We expected that the students would judge the rays' confluence in consistency with the learned thought model, both with respect to the frequency distribution and the judgement average. Our experiments provide evidence for these assumptions but it also became clear that the priming effect of the thought models suggested in the learning examples is not determinative. Though participants always used a variety of thought models, transfer of learning could evoke a shift towards the use of the thought model, which is consistent with the way confluence was portrayed in the learning examples. We also found support for our assumption that the representational image type of rays affects the representation of the confluence situation so that its impact is seen as more severe by the students in the diverging ray image condition.

Although our examples are spatial – i.e. they presuppose the organisation of things in a mentally represented physical space - the crucial differences between them lie in representational contents, namely with such concepts as divergence, confluence, addition, balancing and distribution-based representation. It is the contents referred to by these notions, which must be integrated with spatial information. On the ground of spatial concepts alone, we could hardly find an exhaustive interpretation of the constructed mental representations.

Our results demonstrate that the content of thought models have a decisive effect on a peoples' way of thinking. This implies that content-based analysis is essential. Schema-based analysis and theoretical language would only provide us with information about structural similarities between initial learning examples and the transfer task, but could not explain the substantial variation caused by the content of the thought models. Contents are the essence of representation, and fundamental in explaining the nature of transfer (Saariluoma, 1997, 2003).

Another important point typical of the content-based approach is the necessity to explain individual differences. The distributions in our experiments illustrate how people differ within experimental conditions as to which thought model to use. Some models are very popular while others are less frequently used. Nevertheless, from a content-based point of view all of them are relevant, because psychological analysis of thinking should equally focus on averages of groups as on the differences and variations between them. Naturally, to explain the differences one must have a clear idea about the mental contents of the various groups by identifying similar types of thinking and behaviour.

Apperception is a key concept in content-based theory language, and its link to transfer is central to our argument. In apperception thought models with specific contents are employed to sample stimulus information and to integrate it with information in the memory. The content of information provides us with an essential explanatory framework. In transfer situations, people construct and learn to use new thought models and apply old ones. Individuals differ from each other with regards to the thought models they use in apperceiving presented tasks, and for this reason we obtain very different types of assessments. Some of them may be appropriate and some not. In the first case, we speak about positive transfer and in the latter about negative or failure.

In transfer research, one must be able to state why and how the transfer-related experience and reasoning of person A is different from person B and why person A shows a positive transfer in this situation but not in another. In order to do this, traditional transfer paradigms have referred to either different schemata or the use of different procedural elements in the transfer process (e.g. Singley & Anderson, 1989). Here we argue that eventually the real mental essence of these differences can best be found in information contents.

In sum: The outcomes of our experiments shed more light onto the content-based analysis of apperception. Firstly, there are psychologically relevant questions that can best be explained on the ground of mental contents. Secondly, the ultimate differences among mental representations and respective human thoughts lie in their contents. Thirdly, understanding the construction of mental representation presupposes analysis of their content elements. Fourthly, it is vital to analyse the distributions of different content-elements among individuals instead of searching for an average. The latter approach may easily lead into the loss of essential information. And finally, investigating representational contents as mental contents enables us to ask and answer new types of psychologically relevant questions, as well as to discover new aspects in the old problems.

Our conclusions are ultimately of metascientifical value. It is a common belief that the essence of psychological progress lies in falsifying, i.e., in testing propositions. However, if this objective is pursued too narrow-mindedly it is easily possible to forget that all propositions are constructed from theoretical and empirical concepts, and that for this reason they cannot express what is not already expressible by those concepts. Consequently, if the latter do not have sufficient *power of expression*, one cannot use them to investigate questions that are beyond their scope (Saariluoma, 1997). For instance, one cannot study infinitely small differences between numbers by use of the concept of the natural number, neither can one discuss about the contents of thoughts by applying a language based on the concept of mental capacity or format (Saariluoma, 1997). Hence, to be able to discuss psychological issues in which information contents in mental representations are the relevant explanatory framework, one must develop a suitable methodology and theoretical language.

References

Allard, F. (2001). *Information Processing in Human Perceptual Motor Performance.* Waterloo, Ontario: University of Waterloo, Department of Kinesiology.

Brentano, F. (1874). *Psychologie vom empirischen Standpunkt. I & II.* Leipzig: Duncker & Humblot.

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

Ellis, H. C. (1965). *The transfer of learning.* New York: The Macmillan Company.

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.

Kant, I. (1781/1956). *Kritik der reinen Vernunft.* Darmstadt: Insel.

Leibniz, G. (1704/1989). *New essays on human understanding.* Cambridge: Cambridge University Press.

Neisser, U. (1976). *Cognition and Reality*. San Francisco: W.H. Freeman.

Reed, S. K. (1993). A schema-based theory of transfer. In D. K. Dettermann & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 39-67). Stamford, CT: Ablex Publishing Corp.

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). *Ajattelu työelämässä* [Thinking in work life]. Porvoo: WSOY.

Saariluoma, P. (2003). Apperception, content-based psychology and design. In U. Lindemann (Ed.), *Human behaviours 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. (2003). Stumbling blocks in novice building design. *Journal of Architectural and Planning Research*, *20*(4), 244-254.

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.

Stevens, S. S. (1963). *Handbook of Experimental Psychology*. John Wiley and Sons, Inc., New York.

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

Woodworth, R. S. (1938). *Experimental Psychology.* New York: Holt.

Wundt, W. (1896/1920). *Grundriss der Psychologie.* Stuttgart: Körner.

# Footnotes

[1] In the questionnaire of the distribution-based condition (Experiment 3), instead of the single *small* ray option, the confluence was judged in relation to a single *big* ray (i.e. a four times thicker line was displayed).

[2] The proportion "1 small ray/4 small rays" was computed as one fourth of the proportion "1 big (i.e. four times more intense) ray/4 small rays". As 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).

*Figure 1*

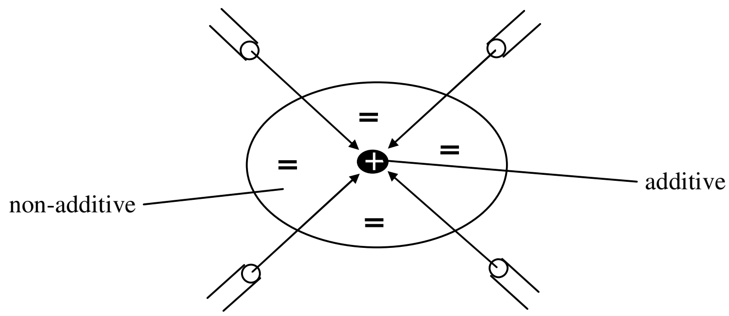


*Figure 2*

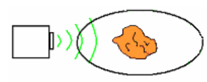
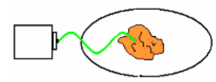
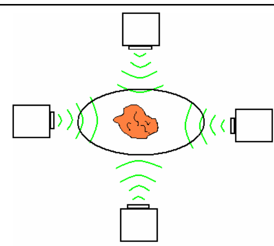
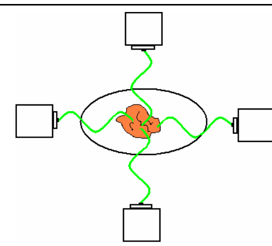
| | Ray image condition | |
|---|---|---|
| | **Diverging** | **Compact** |
| **Single small ray** |  |  |
| **Confluence of four small rays** |  |  |

*Figure 3*

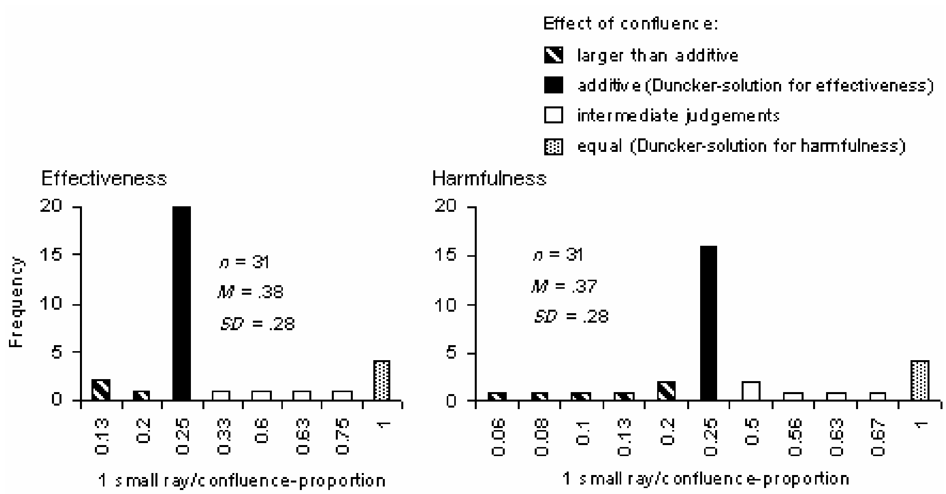

Effect of confluence:
larger than additive
additive (Duncker-solution for effectiveness)
intermediate judgements
equal (Duncker-solution for harmfulness)
Effectiveness
Frequency
20
15
10
5
0
n = 31
M = .38
SD = .28
0.13
0.2
0.25
0.33
0.6
0.63
0.75
1
1 small ray/confluence-proportion
Harmfulness
20
15
10
5
0
n = 31
M = .37
SD = .28
0.06
0.08
0.1
0.13
0.2
0.25
0.5
0.56
0.63
0.67
1
1 small ray/confluence-proportion

*Figure 4*

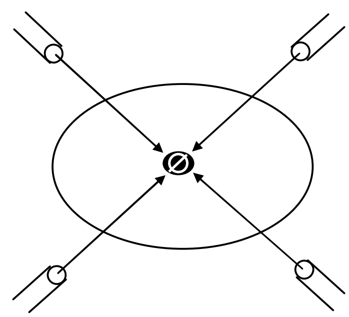


*Figure 5*

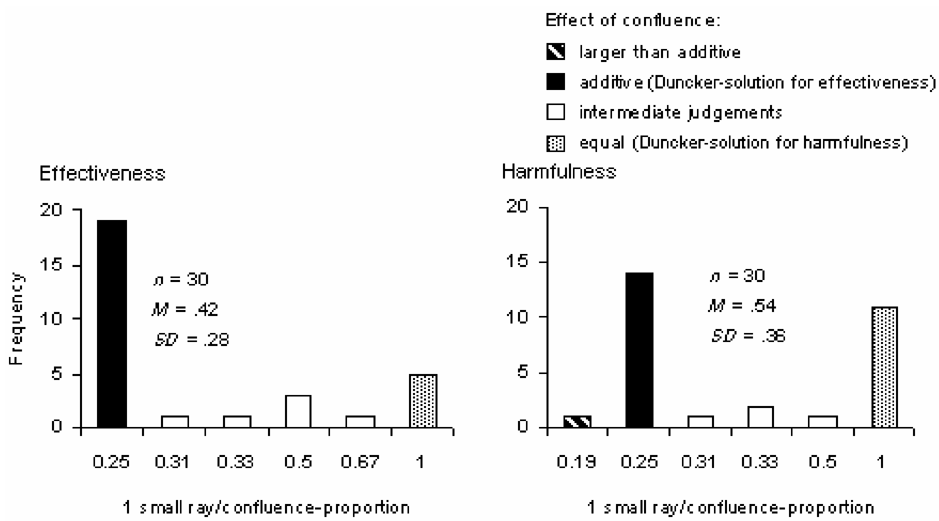

Effect of confluence:
larger than additive
additive (Duncker-solution for effectiveness)
intermediate judgements
equal (Duncker-solution for harmfulness)
Effectiveness
Frequency
n = 30
M = .42
SD = .28
0.25 0.31 0.33 0.5 0.67 1
1 small ray/confluence-proportion
Harmfulness
n = 30
M = .54
SD = .36
0.19 0.25 0.31 0.33 0.5 1
1 small ray/confluence-proportion

*Figure 6*

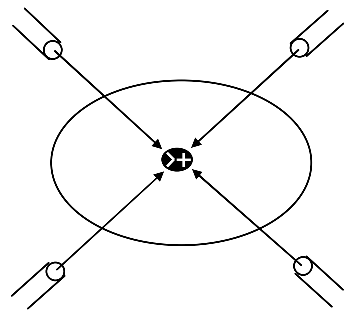


*Figure 7*

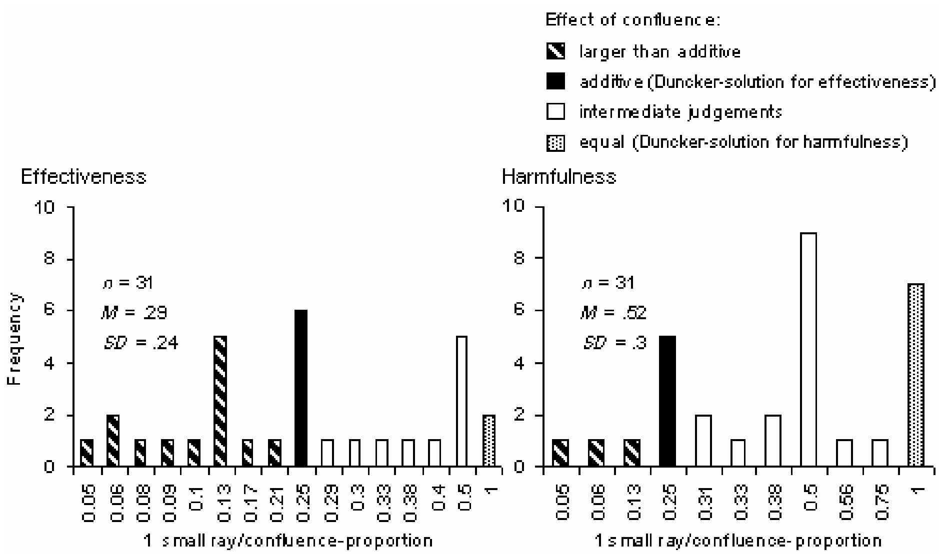
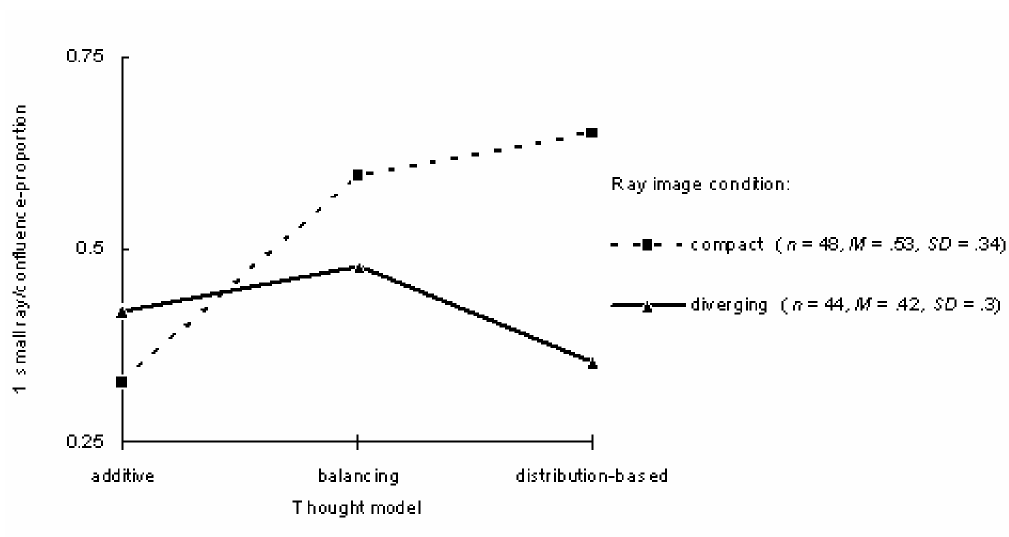

Effect of confluence:
larger than additive
additive (Duncker-solution for effectiveness)
intermediate judgements
equal (Duncker-solution for harmfulness)
Effectiveness
Harmfulness
Frequency
n = 31
M = .29
SD = .24
n = 31
M = .52
SD = .3
0.05 0.06 0.08 0.09 0.1 0.13 0.17 0.21 0.25 0.29 0.3 0.33 0.38 0.4 0.5 1
0.05 0.06 0.13 0.25 0.31 0.33 0.38 0.5 0.56 0.75 1
1 small ray/confluence-proportion
1 small ray/confluence-proportion

*Figure 8*


0.75
0.5
0.25
1 small ray/confluence-proportion
additive
balancing
distribution-based
Thought model
Ray image condition:
compact (n = 48, M = .53, SD = .34)
diverging (n = 44, M = .42, SD = .3)

# Figure Captions

*Figure 1:* The additive and non-additive thought models presumed in Duncker's (1935) ideal representation of a confluence's effectiveness in destroying the tumour and harmfulness to the surrounding tissue respectively. (The ray-beams, the tumour, and the surrounding tissue themselves are displayed in a similar way as Duncker did.)

*Figure 2:* Non-confluence vs. confluence comparison and ray image[1] conditions used in the transfer task of the two questionnaire versions.

*Figure 3:* Assessment of the relative effectiveness and harmfulness of one small ray compared to four confluencing small rays categorized according to the differences in the underlying judgment model.

*Figure 4:* Application of the balancing thought model to the representation of a rays confluence effectiveness for destroying the tumour.

*Figure 5:* Assessment of the relative effectiveness and harmfulness of one small ray compared to four confluencing small rays categorised according to the differences in the underlying judgement model.

*Figure 6:* The thought model where confluence effects exceed the summation of single effects applied to the representation of ray confluence's effectiveness for destroying the tumour.

*Figure 7:* Assessment of the relative effectiveness and harmfulness of one small ray compared to four confluencing small rays categorised according to the differences in the underlying judgement model.

*Figure 8:* Average harmfulness judgements across the three experiments (Experiment 1: additive thought model; Experiment 2: balancing thought model; Experiment 3: distribution-based thought model) split by ray image condition (compact vs. diverging).