VOLUME 58
NUMBER 5

WHOLE NO. 270
1945

Psychological Monographs

JOHN F. DASHIELL, *Editor*

On Problem-Solving

By
KARL DUNCKER
Translated By
LYNNE S. LEES

Published by
THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

Publications Office
MASSACHUSETTS AND NEBRASKA AVENUES, N.W., WASHINGTON 16, D.C.

INTRODUCTION

It is to be regretted that few psychologists take an active part in the investigation of thinking. In modest forms thinking pervades, and to a degree rules, all activities of a human being. Why, then, are we so little concerned with the study of thought processes? It seems that there are two reasons for this attitude. In the first place, many believe that thinking is not a subject matter with any specific characteristics of its own. According to this view further advance in the investigation of associations, of habits, sets, and the like will gradually reduce thought to a mere complication of such simpler facts. In the second place, we like our quantitative experimental procedures so much that a subject matter to which these procedures cannot easily be applied has little chance to arouse our interest. There are serious doubts whether in the field of thought much can be achieved by measuring.

The first of these reasons can obviously not be accepted. When a theoretical hope opens our eyes to new possibilities of research this positive influence of theory is of course entirely desirable. But no theoretical expectation, however strongly it may be ingrained in a given scientific attitude, can be allowed to exclude a subject matter from impartial inspection. How, indeed, will theoretical expectations ever be tested if we provide no independent knowledge of the facts to which these expectations refer and with which they must be confronted in a test?

I do not think that the second reason has greater force than the first. To be sure, we do like to deal with topics for which a ready-made and elegant technique of investigation is available. To penetrate into a field in which standardized procedures are of little avail is not a task by which we feel attracted. But, surely, our personal weaknesses ought not to masquerade as good reasons. If given methods do not fit an unfamiliar subject matter we are under obligation to devise others which give us at least a first acquaintance with the new territory. Precisely this is being done in social psychology, another field in which experimental procedures have not proved particularly successful.

It will be objected that any new endeavor which makes little use of quantitative techniques must necessarily lack the precision of science in a strict sense. Let us not be impressed by this argument. It is familiar to all students of human nature. When an action is not to our taste we tend at once to discover that results will not live up to a certain ideal, and that therefore we ought not to act at all. But in any branch of knowledge, just as in politics, in economic and in social issues, our conduct will almost invariably be defined as right only by its *direction;* we have little commerce with perfection. To point to the imperfections of first steps is under these circumstances almost a symptom of inertia. How do the critics plan ever to conquer difficult new ground if the absence of paved roads seems to them sufficient reason for keeping out entirely?

I wish to express my gratitude to Professor E. G. Boring who was kind enough to make the library facilities of the Department of Psychology at Harvard University available to me. Without this help I could not have done the present translation. I am also indebted to Professor W. Köhler who compared the translation with the original, and made suggestions for the improvement of the English version.

LYNNE S. LEES

In view of so much hesitation any courageous attempt to throw light upon the nature of thinking must be welcomed. It is only such attempts which can gradually make us familiar with the specific problems and with the technical possibilities of this field. In 1935 Karl Duncker published his book *Zur Psychologie des produktiven Denkens* which, it seems to me, has penetrated farther into its subject than most other enterprises with a similar program. But few psychologists and hardly any philosophers in this country are acquainted with this important investigation. Moreover, for years the original has been virtually unobtainable. It is therefore most fortunate that after Duncker's death one of his students decided to prepare an English translation. Those who know the German text will be able to realize the difficulty of Mrs. Lees' task; the skill with which she has overcome this difficulty seems to me admirable.

This is not an easy book. The author did not write it when the hard labors of research lay far behind him so that he could serenely look upon his findings from a distance. Rather, every page seems to show him in the midst of his untiring struggle for clearness. Thus to our interest in his analysis as such is added a more particular attraction which few books have.

Duncker's work is not easy for a further reason. He hated any compromise with vague terms which give an appearance of knowledge while they actually hide problems. He simply would not let go until he knew the very anatomy of a concept. I also have the impression that once he had started on his journey he was not satisfied until he had visited all provinces of thought. Thus in some chapters we find the author far within philosophy because the ramifications of insight would not become clear without that transgression. In other chapters he studies the obstacles to productive thought rather than its course when it succeeds. There are sections from which all teachers of mathematics could greatly profit, and many pages in which the investigation of thought elucidates problems of perception.

Nobody who reads this book will be able to deny that fruitful research is possible in the realm of thought. Moreover, in several points Duncker confirms the work of others. Clearly, observation in this field is largely independent of theoretical preconceptions. To one conclusion or another the reader with a training in psychology may be inclined to object. If that happens he will also like to prove his point—and will thus find himself working on thought processes. It *is* time that we turn to essentials.

In Duncker himself this was the greatest intellectual virtue: He was forever impatient of little things and happy only when he found a way that led to fundamentals. The best we can do in remembering our friend is to give his work as an example which others may follow.

WOLFGANG KÖHLER

FOREWORD

To study productive thinking where it is most conspicuous in great achievements is certainly a temptation, and without a doubt, important information about the genesis of productive thought could be found in biographical material. But although a thunderstorm is the most striking example of electrical discharge, its laws are better investigated in little sparks within the laboratory. To study in simple, convenient forms what is complicated and difficult of access is the method of experimental science; to lose in this simplification just the essential aspects, is its notorious danger. Experimental psychology, more than all other sciences, is continually faced with this danger. I hope to have succeeded in simplifying my subject without altering its essential nature.

The restriction of the investigation to practical and mathematical problems in thinking is entirely intentional. This restriction was adopted because such material is more accessible, more suitable for experimentation. I believe, however—and the reader may agree with me at the end of the book—that essential features of problem-solving are independent of the specific thought-material. But there is still another restriction for which I must account to the reader. In my analysis, all facts are intentionally omitted which have no immediate significance for the *problem of finding*, of εὑρίσκειν—be they ever so essential parts of the complete psychological inventory of a solution-process. According to the modern psychology of needs and affective states, every thought-problem which the subject actually makes his own has at its base, as its source of energy, a system of tensions of the type of needs (Lewin). In the course of a solution-process, this tension-system undergoes all kinds of alterations until, where possible, it finds complete release in a final solution. The fate of this tension-system, as shown in rising and subsiding interest, in experiences of success and of failure, in digressions, substitute behavior, resignation and anger, or in pride of performance and elevated level of aspiration—all these and similar aspects of a solution-process have with a certain ruthless consistency been left untouched in the following analysis. For, they are irrelevant to the specific question of finding: In what way can a meaningful solution be found at all? And in addition, there are already many excellent papers on just these things, while the problem of finding has received scant attention in modern psychology.

Although everything in this investigation is, so to speak, under the command of this problem of finding, much may also be profitably taken from the viewpoint of quite different problems. Such is above all the discussion of *insight* and of *evidence* contained in Chapters IV and V, which may claim a certain theoretical independence. It represents an attempt partially to solve, from a new approach, the old epistemological problem of the recognition of "intrinsic necessity". This discussion continues a philosophical development which is to a degree characterized by the names of Hume, Kant, Husserl and Wertheimer.

Here I should like to express my gratitude to the Abraham Lincoln Foundation (U.S.A.), which in the year 1929, through its financial assistance, made possible comprehensive studies on the problem of mathematical talent, some of which are reported in Chapter III and chiefly in Chapter VIII.

This book is gratefully dedicated to my teachers Wolgang Köhler and Max Wertheimer.

KARL DUNCKER

Berlin, February 1935

CONTENTS

PART I

The Structure and Dynamics of Problem-Solving Processes

PART II

Insight, Learning and Simple Finding

PART III

Fixedness of Thought-Material

Part One

THE STRUCTURE AND DYNAMICS OF PROBLEM-SOLVING PROCESSES

Chapter I

The Solution of Practical Problems (I)

1. INTRODUCTION AND FORMULATION OF THE PROBLEM

A PROBLEM arises when a living creature has a goal but does not know how this goal is to be reached. Whenever one cannot go from the given situation to the desired situation simply by action, then there has to be recourse to thinking. (By action we here understand the performance of obvious operations.) Such thinking has the task of devising some action which may mediate between the existing and the desired situations. Thus the 'solution' of a practical problem must fulfill two demands: in the first place, its realization[1] must bring about the goal situation, and in the second place one must be able to arrive at it from the given situation simply through action.

The practical problem whose solution was experimentally studied in greatest detail runs as follows: Given a human being with an inoperable stomach tumor, and rays which destroy organic tissue at sufficient intensity, by what procedure can one free him of the tumor by these rays and at the same time avoid destroying the healthy tissue which surrounds it?

Such practical problems, in which one asks, "How shall I attain . . .?", are related to certain theoretical problems, in which the question is, "How, by what means, shall I comprehend . . .?" In the former case, a problem situation arises through the fact that a goal has no direct connection with the given reality; in the latter case—in theoretical problems—it arises through the fact that a proposition has no direct connection with what is given in the premises. As example in the latter field, let us take again the problem with which I experimented in greatest detail: Why is it that all six-place numbers of the type *abcabc,* for example 276276, are divisible by thirteen?

It is common to both types of problems that one seeks the ground for an anticipated consequence; in practical problems, the actual ground is sought; in theoretical problems, the logical ground.[2]

In the present investigation the question is: *How does the solution arise from the problem situation? In what ways is the solution of a problem attained?*

2. EXPERIMENTAL PROCEDURE

The experiments proceeded as follows:

[1] [Translator's note: "Realization" is used in the sense of "making real", of "actualization". The terms "embodiment" and "to embody" are used in a closely related sense, which will be clear in context. In the following, all notes of the translator will be given in parentheses. Such notes will add the German terms of the original where entirely satisfactory English terms do not seem to exist.]

[2] Other types of theoretical problems, such as: "What is the essential nature of, or the law of . . . ?" or "How are . . . related to each other?", are not investigated here.

The subjects (Ss), who were mostly students of universities or of colleges, were given various thinking problems, with the request that they think aloud. This instruction, *"Think aloud"*, is not identical with the instruction to introspect which has been common in experiments on thought-processes. While the introspecter makes himself as thinking the object of his attention, the subject who is thinking aloud remains immediately directed to the problem, so to speak allowing his activity to become verbal. When someone, while thinking, says to himself , "One ought to see if this isn't . . .", or, "It would be nice if one could show that . . .", one would hardly call this introspection; yet in such remarks something is revealed which we shall later deal with under the name of 'development of the problem'. The subject (S) was emphatically warned not to leave unspoken even the most fleeting or foolish idea. He was told that where he did not feel completely informed, he might freely question the experimenter, but that no previous specialized knowledge was necessary to solve the problems.

3. A PROTOCOL OF THE RADIATION PROBLEM

Let us begin with the radiation problem (p. 1). Usually the schematic sketch shown in Fig. 1 was given with the problem. Thus, it was added, somebody had visualized the situation to begin with (cross-section through the body with the tumor in the middle and the radiation apparatus on the left); but obviously this would not do.

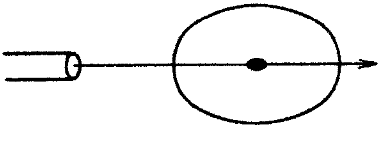


FIG. 1

From my records I choose that of a solution-process which was particularly rich in typical hunches and therefore also especially long and involved. The average process vacillated less and could be left to run its own course with considerably less guidance.[3]

Protocol

1. Send rays through the esophagus.
2. Desensitize the healthy tissues by means of a chemical injection.
3. Expose the tumor by operating.
4. One ought to decrease the intensity of the rays on their way; for example—would this work?—turn the rays on at full strength only after the tumor has been reached. (Experimenter: False analogy; no injection is in question.)
5. One should swallow something inorganic (which would not allow passage of the rays) to protect the healthy stomach-walls. (E: It is not merely the stomach-walls which are to be protected.)
6. Either the rays must enter the body or the tumor must come out. Perhaps one could alter the location of the tumor—but how? Through pressure? No.
7. Introduce a cannula.—(E: What, in general, does one do when, with any agent, one wishes to produce in a specific place an effect which he wishes to avoid on the way to that place?)
8. (Reply:) One neutralises the effect on the way. But that is what I have been attempting all the time.
9. Move the tumor toward the exterior. (Compare 6.) (The E repeats the problem and emphasizes, ". . . which destroy *at sufficient intensity*".)
10. The intensity ought to be variable. (Compare 4.)
11. Adaptation of the healthy tissues by previous weak application of the rays. (E: How can it be brought about that the rays destroy only the region of the tumor?)

[3] Compare the pertinent protocols in my earlier and theoretically much less developed paper, "A qualitative study of productive thinking," Ped. Sem., 1926, v. 33.

12. (Reply:) I see no more than two possibilities: either to protect the body or to make the rays harmless. (E: How could one decrease the intensity of the rays en route? [Compare 4.])

13. (Reply:) Somehow divert . . . diffuse rays . . . disperse . . . stop! Send a broad and weak bundle of rays through a lens in such a way that the tumor lies at the local point and thus receives intensive radiation.[4] (Total duration about half an hour.)

4. IMPRACTICABLE 'SOLUTIONS'

In the protocol given above, we can discern immediately that the whole process, from the original setting of the problem to the final solution, appears as a series of more or less concrete proposals. Of course, only the last one, or at least its principle, is practicable. All those preceding are in some respect inadequate to the problem, and therefore the process of solution cannot stop there. But however primitive they may be, this one thing is certain, that they cannot be discussed in terms of meaningless, blind, trial-and-error reactions. Let us take for an example the first proposal: "Send rays through the esophagus". Its clear meaning is that the rays should be guided into the stomach by some passage free from tissue. The basis of this proposal is, however, obviously an incorrect representation of the situation inasmuch as the rays are regarded as a sort of fluid, or the esophagus as offering a perfectly straight approach to the stomach, etc. Nevertheless, within the limits of this simplified concept of the situation, the proposal would actually fulfill the demands of the problem. It is therefore genuinely the solution of a problem, although not of the one which was actually presented. With the other proposals, the situation is about the same. The second presupposes that a means—for example, a chemical means—exists for making organic tissue insensitive to the rays. If such a means existed, then everything would be in order, and the solution-process would have already come to an end. The fourth proposal—that the rays be turned on at full strength only when the tumor has been reached—shows again very clearly its derivation from a false analogy, perhaps that of a syringe which is set in operation only when it has been introduced into the object. The sixth suggestion, finally, treats the body too much as analogous to a rubber ball, which can be deformed without injury. In short, it is evident that such proposals are anything but completely meaningless associations. Merely in the factual situation, they are wrecked on certain components of the situation not yet known or not yet considered by the subject.

Occasionally it is not so much the situation as the demand, whose distortion or simplification makes the proposal practically useless. In the case of the third suggestion, for example ("expose the tumor by operating"), the real reason why radiation was introduced seems to have escaped the subject. An operation is exactly what should be avoided. Similarly in the fifth proposal, the fact is forgotten that not only the healthy stomach-walls must be protected but also all parts of the healthy body which have to be penetrated by the rays.

A remark on principle may here be in order. The psychologist who is investigating, not a store of knowledge, but the genesis of a solution, is not interested

[4] This solution is closely related to the 'best' solution: *crossing of several weak bundles of rays at the tumor,* so that the intensity necessary for destruction is attained only here. Incidentally, it is quite true that the rays in question are not deflected by ordinary lenses; but this fact is of no consequence from the viewpoint of the psychology of thinking. See 4 below.

primarily in whether a proposal is actually practicable, but only in whether it is formally practicable, that is, practicable in the framework of the subject's given premises. If in planning a project an engineer relies on incorrect formulae or on non-existent material, his project can nevertheless follow from the false premises as intelligently as another from correct premises. One can be 'psychological equivalent' to the other. In short, we are interested in knowing how a solution develops out of the system of its subjective premises, and how it is fitted to this system.

5. CLASSIFICATION OF PROPOSALS

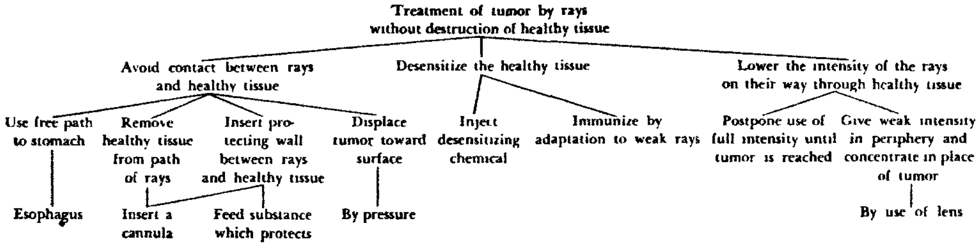
If one compares the various tentative solutions in the protocol with one another, they fall naturally into certain groups. Proposals 1, 3, 5, 6, 7 and 9 have clearly in common the attempt to *avoid contact between the rays and the healthy tissue*. This goal is attained in quite different ways: in 1 by re-directing the rays over a path naturally free from tissue; in 3 by the removal of the healthy tissue from the original path of the rays by operation; in 5 by interposing a protective wall (which may already have been tacitly implied in 1 and 3); in 6 by translocating the tumor towards the exterior; and in 7, finally, by a combination of 3 and 5. In proposals 2 and 11, the problem is quite differently attacked: the accompanying destruction of healthy tissue is here to be avoided by the *desensitizing or immunizing of this tissue*. A third method is used in 4, perhaps in 8, in 10 and 13: *the reduction of radiation intensity on the way*. As one can see, the process of solution shifts noticeably back and forth between these three methods of approach.

In the interests of clarity, the relationships described are presented graphically on the next page.

6. FUNCTIONAL VALUE AND UNDERSTANDING

In this classification, the tentative solutions are grouped according to the manner in which they try to solve the problem, i.e., according to their "by-means-of-which", their "functional value". Consider the proposal to send rays through the esophagus. The S says nothing at all about avoiding contact, or about a free passage. Nevertheless, the solution-character of the esophagus in this context is due to no other characteristic than that of being a tissue-free path to the stomach. It functions as the embodiment solely of this property (not of the property of being a muscular pipe, or of lying behind the windpipe, or the like). In short, in the context of this problem, the "by-means-of-which", the "functional value" of the esophagus is: a free path to the stomach. The proposals: "direct the rays by a natural approach", "expose by operation", "translocate the tumor toward the exterior", "protective wall", and "cannula" all embody the functional value: no contact between rays and healthy tissue. The functional value of the solution, "concentration of diffuse rays in the tumor", is the characteristic: "less intensity on the way, great intensity in the tumor". The functional value of the lens is the quality: "medium to concentrate rays", and so forth.

The functional value of a solution is indispensable for the understanding of its being a solution. It is exactly what is called the sense, the principle or the point of the solution. The subordinated, more specialized characteristics and properties of a solution embody this principle, apply it to the particular circum-



stances of the situation. For example, the esophagus is in this way an application of the principle: "free passage to the stomach", to the particular circumstances of the human body. To understand the solution as a solution is just the same as to comprehend the solution as embodying its functional value. When someone is asked, "Why is such-and-such a solution?", he necessarily has recourse to the functional value. In all my experiments, aside from two or three unmistakable exceptions, when the E asked about a proposal: "In what way is this a solution of the problem?", the S responded promptly with a statement of its functional value. (In spontaneous statements of the Ss, the functional value was frequently left unmentioned as being too obvious.)

Incidentally, the realization of its functional value mediates understanding of a solution even where there is nothing but an 'unintelligible' (though sufficiently general) relation between the functional value and the demand which it fulfills. Blowing on a weakly glimmering fire, for example, undoubtedly solves the problem of rekindling the fire because in this way fresh oxygen is supplied. In other words, the increase of the oxygen supply is the immediate functional value of blowing on the fire. But why combination with oxygen produces warmth and flame is ultimately not intelligible. Even if the whole of chemistry should be successfully and without a gap derived from the principles of atomic physics, these principles are not in themselves altogether intelligible. i.e., ultimately they must be "accepted as mere facts". (See details in Chap. IV, especially page 47f.) Thus, intelligibility frequently means no more than participation in, or derivability from, sufficiently elementary and universal causal relationships. But even if these general laws are not in themselves intelligible, reducibility to such general laws actually mediates a certain type of understanding.

To the same degree to which a solution is understood, it can be transposed, which means that under altered conditions it may be changed correspondingly in such a way as to preserve its functional value. For, one can transpose a solution only when one has grasped its functional value, its general principle, i.e., the invariants from which, by introduction of changed conditions, the corresponding variations of the solution follow each time.

An example: When, seen from the standpoint of a spectator, someone makes a detour around some obstacle, and yet acts from his own point of view in terms of nothing but, say, "now three yards to the left, then two yards straight ahead, then to the right . . ."—these properties of the solution would certainly satisfy the concrete circumstances of the special situation here and now. But so long as the person in question has not grasped the functional value, the general structure: "detour around an obstacle", he must necessarily fail when meeting a new obstacle which is differently located and of different shape. For to different obstacles correspond different final forms of the solution; but the structure, "detour around an obstacle", remains always the same. Whoever has grasped this structure is able to transpose a detour properly.

7. MEANINGLESS ERRORS AS A SYMPTOM OF DEFICIENT UNDERSTANDING

A solution conceived without functional understanding often betrays itself through nonsensical errors. A good example is supplied by experiments with another thinking problem.

The problem was worded as follows: "You know what a pendulum is, and that a pendulum plays an important rôle in a clock. Now, in order for a clock to go accurately, the swings of the pendulum must be strictly regular. The duration of a pendulum's swing depends, among other things, on its length, and this of course in turn on the temperature. Warming produces expansion and cooling produces contraction, although to a different degree in different materials. Thus every temperature-change would change the length of the pendulum. But the clock should go with absolute regularity. How can this be brought about? —By the way, the length of a pendulum is defined solely by the shortest distance between the point of suspension and the center of gravity. We are concerned only with this length; for the rest, the pendulum may have any appearance at all."

The customary solution of this pendulum problem in actual practice is reproduced in Figure 2. At first this solution

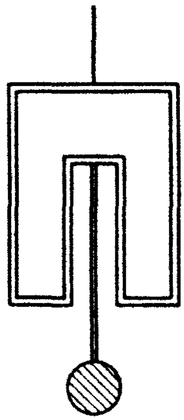


FIG. 2

will be entirely 'unintelligible' to many a reader.

Let him watch now what takes place when the solution suddenly becomes clear to him. Its functional value is that every expansion in one direction is compensated by an equally great expansion in the opposite direction.

The bars *a* and *a'* (see Fig. 3) can expand only downwards; *b* and *b'*, on the other hand, only upwards, since they are fastened below. The bars *b* and *b'* are meant to raise the strip of metal to which *c* is fastened by exactly as much as *a* and *c* together expand downwards. To this end, *b* and *b'* must of course be constructed of a material with a greater coefficient of expansion than *a* and *a'* and *c*.

Only when Figure 3 is grasped as the

embodiment of this functional value, is it understood as the solution.

Among the many Ss to whom I gave the pendulum problem, there were two who were already vaguely familiar with a pendulum-model, and simply reconstructed it from memory. One was fortunate and did it correctly, while the

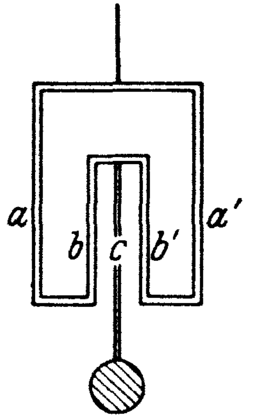


FIG. 3

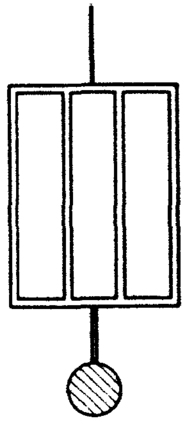


FIG. 4

other drew "just four or five bars like this, from which the weight hung below". (Fig. 4) It is evident that this is a completely meaningless construction, despite all external resemblance to Fig. 3, and devoid of any functional understanding (as the S clearly realized and expressed himself). Compare with this the solutions of the problem contained in Figure 5, a-g, which, in spite of all external differences, embody the identical functional value and at the same time represent completely new constructions.

In all of them there is compensation in the sense of Figure 3; thus we deal with appropriate transpositions of Figure 3. It is worth mentioning that one S drew the model of Figure 5a, believing that it was the compensation-pendulum dimly familiar from experience. Here it is clear that the reconstruction can have taken place only via the common functional value. Nothing in their form is common to the two pendulums.

"Good" and "stupid" errors in Köhler's[5] sense can be clearly distinguished as follows: In the case of good, intelligent errors, at least the general functional value of the situation is correctly outlined, only the specific manner of its realization is not adequate. For example, an ape stands a box on its corner under the goal object, which hangs high above, because in this way the box comes closer—to be sure, at the price of its stability. In the case of stupid errors, on the other hand, the outward form of an earlier, or an imitated solution is blindly reproduced without functional understanding. For example, an ape jumps into the air from a box—but the goal object is hanging at quite a different spot.

8. THE PROCESS OF SOLUTION AS DEVELOPMENT OF THE PROBLEM

It may already have become clear that the relationship between superordinate and subordinate properties of a solution has *genetic* significance. *The final form*

[5] Köhler, (20), p. 194, 217.

of an individual solution is, in general, not reached by a single step from the original setting of the problem; on the contrary, the principle, the functional value of the solution, typically arises first, and the final form of the solution in question develops only as this principle becomes successively more and more concrete. In other words, the general or "essential" properties of a solution genetically precede the specific properties; the latter are developed out of the former. The classification given on page 5 presents, thus, a sort of *"family tree"* of the solution of the radiation problem.

The finding of a general property of a solution means each time a *reformulation of the original problem.* Consider for example the fourth proposal in the protocol above. Here it is clearly evident that at first there exists only the very general functional value of the solution: "one must decrease the intensity of the radiation on the way". But the decisive re-formulation of the original problem is thereby accomplished. No longer, as at the beginning, does the S seek simply a "means to apply rays to the tumor without also destroying healthy tissue", but already—over and above this—a means to decrease the intensity of the radiation on the way. The formulation of the problem has thus been made sharper, more specific—and the proposal not to turn the rays on at full strength until the tumor has been reached, although certainly wrong, arises only as a solution of this new, re-formulated problem. From this same reformulation of the problem there arises, at the end of the whole process, the practicable solution, "concentration of diffuse rays in the tumor". With the other proposals in the protocol, the case is similar: the solution-properties found at first, the functional values, *always serve as productive reformulations of the original problem.*[6]

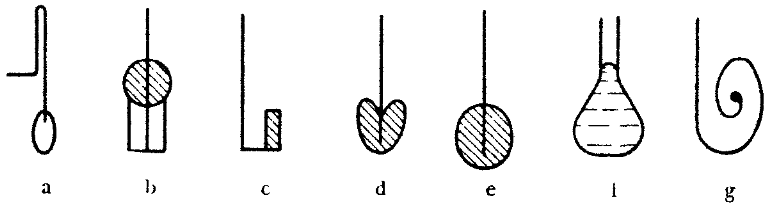


FIG. 5, a-g

We can accordingly describe a process

[6] O. Selz, in his experiments on problem-solving, has already found similar transformations of problems. (Selz, (36), p. 87; (37), p. 41.) Selz defines transformation of the task as the "exchange of the original goal for a more specific one," or, more explicitly, "the substitution for the task of another task, through whose solution the original problem is also to be solved." I give two examples from his experimental repertoire. The first runs as follows: "Give two species of electoral franchise." The first thing which occurs to one S is "that he is familiar with several classifications of electoral franchise; and now he sets himself the specific task of reproducing one of these classifications." Another S realizes "that there exists a contrast more extreme than that in the electoral franchises of North and South Germany; and now he tries to define more closely that extreme contrast; thus the more familiar definition by location again precedes the less familiar definition by content." (36), p. 65; for the protocol itself, see p. 40.)

A second example: In connection with the task of finding "a concept co-ordinated with 'railway platform,'" a transformation of the task attained by many Ss is to produce "another part of the concrete spatial whole which includes 'railway platform'" (which leads, for example, to the solution, 'railroad track'); cf. (37), p. 142.

Selz calls transformations which are severally ordered to whole types of problems, "methods of solution".

of solution either as development of the solution or as development of the problem. Every solution-principle found in the process, which is itself not yet ripe for concrete realization, and which therefor fulfills only the first of the two demands given on page 1, functions from then on as reformulation, as sharpening of the original setting of the problem. *It is therefore meaningful to say that what is really done in any solution of problems consists in formulating the problem more productively.*

To sum up: *The final form of a solution is typically attained by way of mediating phases of the process, of which each one, in retrospect, possesses the character of a solution, and, in prospect, that of a problem.*

At the same time it is evident that, generally speaking, a process of solution penetrates only by degrees into the more specific circumstances and possibilities of the given situation. In the phase, "avoiding contact between rays and healthy tissue", for example, there is still very little reference to the concrete individuality of the situation. The rays function for the time being as "active agent", the tumor as "the place to be influenced", and the healthy tissue as "surrounding region which must be protected". In the next phase, "redirection of the rays over a tissue-free path to the stomach", at least the possibility of such a displacement of the rays is already made use of. In the search for a free pathway, the situation is then subjected to an even more precise inspection; as a consequence, such a specific component of the situation as the esophagus enters the solution-process and is used in a sensible manner.

To widen our horizon, let us here demonstrate with a mathematical example how a solution-process typically arrives at the final solution by way of mediation problem- or solution-phases. The original problem is to prove that there is an infinite number of prime numbers (to find "something from which follows that there exists . . ."). A step which is quite decisive, although subjectively hardly noted, consists in the solution-phase: "I must prove that for any prime number p there exists a greater one". This reformulation of the problem sounds quite banal and insignificant. Nevertheless I had Ss who never hit on it. And without this step, the final solution cannot be reached.[7]

A further solution-phase would run as follows: "To prove the existence of such a prime number, I must try to construct it." With one of my Ss, I could follow clearly the way in which, to this phase, a further one attached itself as a mere explication: "One must therefore construct a number greater than p which cannot be represented as a product". From here on, clearly directed to "avoiding a product", the S proceeded to construct the product of all numbers from 1 to p and to add 1—incidentally, without having realized that the resultant nummer need not be itself a prime number, but may merely contain the desired number as a fraction of itself.

9. IMPLICIT SOLUTION-PHASES

Not all phases of the various solution-processes are given in a family tree of the

[7] The solution consists in the construction of the product of all prime numbers from 1 to p and adding to it 1. The resultant number is either itself a prime number, or it is a product of prime numbers greater than p. For, with the exception of the special case of 1, a prime number less than p cannot be contained in a multiple of itself increased by 1 without a remainder. Thus in any case, a prime number greater than p exists. (Q.E.D.)

kind graphically represented on page 5: rather, only the more prominent and relatively independent among them are given. Aside from these, there exist phases which are not explicit enough and, above all, too, banal ever to appear in a protocol. In the case of the radiation problem, for instance, it is clear to all Ss from the start that, in any case, to find a solution, something must be done with the actual circumstances concerned, with the rays and the body. As modern Europeans, they do not think of looking for suitable magic formulae; nor would they anticipate that some change in another place would lead to a solution. Similarly in the case of the prime numbers problem, from the beginning there is no doubt that the solution is to be sought in the province of numbers, and not, for example, in the province of physical processes. In short, from the very first, the deliberating and searching is always confined to a province which is relatively narrow as to space and content. Thus preparation is made for the more discrete phases of a solution by certain *approximate regional demarcations,* i.e., by phases in which necessary but not yet sufficient properties of the solution are demanded. Such implicit phases of a solution do not quite fulfill even the first prerequisite of a solution mentioned on page 1.

This is valid not only for thinking, but also for attempts at solution by action (trial and error). When a layman wishes to adjust the spacing between lines on a typewriter, this much at least of the solution is known to him: "I must screw or press somewhere on the machine". He will not knock on the wall, for instance, nor does he anticipate that any change of the given colors would do. In general, one seeks to achieve mechanical effects by mechanical alterations in the critical object.

One more example, this time from animal psychology. Thorndike (39) set his experimental animals (mostly cats) problems of the following type. They had to learn to bring about the opening of their cage doors by a simple mechanical manipulation—unintelligible to them, to be sure, for they could not survey the connections—and so to escape into freedom. Part of the animals had a whole series of different cage problems to solve. In one cage they had to pull on a loop, in another to lift a bar, or press on a knob, etc. Thorndike made the very interesting observation that generally, in the course of the experiments, "the cat's general tendency to claw at loose objects within the box is strengthened and its tendency to squeeze through holes and bite bars is weakened." (p. 48) Further, "its tendency to pay attention to what it is doing gets strengthened . . ." (p. 48) It is evident that even animal 'trial and error' is for the most part already under the confining influence of certain demarcations, which, by the way, are not purely instinctive.

10. INSUFFICIENCY OF A PROTOCOL

The reader has probably received the impression that the discussions of the preceding paragraphs left the data of the protocol a long way behind. In the case of the very first proposal, for instance, that of the esophagus, there was no mention at all of "redirecting over a tissue-free path", or even of "avoiding contact". That some such thing appeared in other protocols in an analogous place naturally proves nothing about the psychological origin of just this individual proposal.

This is the place in which to say

something essential about protocols. One could formulate it thus: A protocol is relatively reliable only for what it positively contains, but not for that which it omits. For even the best-intentioned protocol is only a very scanty record of what actually happens. The reasons for this insufficiency of protocols which are based on spoken thoughts must interest us also as characteristic of a solution-process as such. Mediating phases which lead at once to their concrete final realization, and thus are not separated from the solution by clear phase-boundaries, will often not be explicitly mentioned. They blend too thoroughly with their final solutions. On the other hand, mediating phases which must persist as temporary tasks until they find their final 'application' to the situation have a better chance of being explicitly formulated. Furthermore, many superordinate phases do not appear in the protocol, because the situation does not appear to the S promising enough for them. Therefore they are at once suppressed. In other words, they are too fleeting, too provisional, too tentative, occasionally also too 'foolish', to cross the threshold of the spoken word.

In very many cases, the mediating phases are not mentioned because the S simply does not realize that he has already modified the original demand of the problem. The thing seems to him so self-evident that he does not have at all the feeling of having already taken a step forward.[8] This can go so far that the S deprives himself of freedom of movement to a dangerous degree. By substituting unawares a much narrower problem for the original, he will therefore remain in the framework of this narrower problem, just because he confuses it with the original.[9]

11. "SUGGESTION FROM BELOW"

There exist cases in which the final form of a solution is not reached from above, i.e., not by way of its functional value. This is a commonplace of 'familiar' solutions. If the final solution of a problem is familiar to the S, it certainly need no longer be constructed, but can be reproduced as a whole, as soon as the problem is stated.[10]

More interesting cases exist. We must always remember that a solution has, so to speak, two roots, one in that which is sought and one in that which is given. More precisely, *a solution arises from the claim made on that which is given by that which is sought. But these two components vary greatly in the share they have in the genesis of a solution-phase.* A property of a solution is often very definitely demanded (characterized, hinted at) before it is discovered in what is given; but sometimes it is not. An example from the radiation problem: The esophagus may be discovered because a free path to the stomach is already sought. But it may also happen that, during a relatively vague, planless inspection of what is given in the situation, one 'stumbles on the esophagus'. Then the latter—so to speak, from below—suggests its functional value: "free path to the stomach"; in other words, the concrete realization precedes the functional value. This sort of thing happens not infrequently; for the analysis of the situation is often relatively planless. Nor is this disadvanta-

[8] Such is especially the case with the demarcation of boundaries, cf. pages 9-10.

[9] Cf. pages 25-26, on "Fixation".

[10] This of course does not exclude the possibility that the solution is reproduced along with its functional value and as its realization, and that it is thus *understood.*

geous, when the point is to find new ideas. In mathematical problems, this analysis merely of the given situation, the development of consequences from the given data, plays an especially large rôle —as we shall see in Chapter III.

One more example of "suggestion from below". An attractive goal object (for example, a banana) lies out of reach before the cage of a chimpanzee. So long as the solution, "to fish for the banana with a stick", is not very familiar, something like a stick must be in the visual field as a suggesting factor.[11] The stick is not yet *sought*—as embodiment of the previously conceived functional value: 'something long and movable"—as it is in later stages; rather it must itself help to suggest this functional value.[12]

The prerequisite for such a suggestion from below is that the 'phase-distance' between what is sought for and what could give the suggestion is not too great.

The following is an example for this influence of the size of the phase distance. Right at the beginning of the radiation problem, the E can speak of "crossing", or can draw a cross, without the S's grasping what that means. (Cf. the solution by crossing a number of weak bundles of rays in the tumor.) If, on the other hand, the S is already of his own accord directed to "decreasing the intensity on the way", he will understand the suggestion sooner than if his thinking is dominated, for example, by the completely different demand for "a free path for the rays". We can formulate the general proposition that a suggestion is the sooner understood or assimilated, the closer it approaches the genealogical line already under development, and, within this line, the nearer it is to the problem-phase then in operation; in short, the more completely it is already anticipated.

This law is a special case of a more general law, which concerns not suggestions in the narrow sense, but the material of thinking in general. Selz formulated this law as "a general law of anticipation" in the following manner: "An operation succeeds the more quickly, the more the schematic anticipation of the solution approaches a complete anticipation." (37, p. 512) We shall have more to do with this law; see Chapter VI.

12. LEARNING FROM MISTAKES (CORRECTIVE PHASES)

As yet we have dealt only with the progress from the superordinate to the subordinate phases (or vice versa), in other words, with progress along a given genealogical line. That this is not the only kind of phase succession is, one should think, sufficiently indicated by the protocol given above. Here the line itself is continually changed, and one way of approach gives way to another. Such a *transition to phases in another line* takes place typically when some tentative solution does not satisfy, or when one makes no further progress in a given direction. *Another* solution, more or less clearly defined, is then looked for. For instance, the first proposal (esophagus) having been recognized as unsatisfactory, quite a radical change in direction takes place. The attempt to avoid contact is completely given up and a

[11] Cf. W. Köhler, (20), p. 37.

[12] This suggestion of the functional value from below is even the rule in problems where a number of objects are offered to begin with, with the instruction to choose from among them an appropriate tool for such and such a purpose. Especially when only few objects are concerned, thinking will tend to proceed by looking things over, i.e., it will test the given objects one after another as to their applicability, and no attempt will be made to conceive the appropriate functional value first. (See Chap. VII on experiments with such problem-situations.)

means to desensitize tissues is sought in its place. In the third proposal, however, the S has already returned to old tactics, although with a new variation. And such shifting back and forth occurs frequently.

It will be realized that, in the transition to phases in another line, the thought-process may range more or less widely. Every such transition involves a return to an earlier phase of the problem; an earlier task is set anew; a new branching off from an old point in the family tree occurs. Sometimes a S returns to the original setting of the problem, sometimes just to the immediately preceding phase. An example for the latter case: From the ingenious proposal, to apply the rays in adequate amounts by rotation of the body around the tumor as a center, a S made a prompt transition to the neighboring proposal: "One could also have the radiation apparatus rotate around the body." Another example: The S who has just realized that the proposal of the esophagus is unsatisfactory may look for another natural approach to the stomach. This would be the most "direct" transition, that is, the transition which retrogresses least. Or, renouncing the natural approach to the stomach, he looks for another method of avoiding contact. Or, again, he looks for an altogether different way to avoid the destruction of healthy tissue. Therewith, everything which can be given up at all would have been given up; a "completely different" solution would have to be sought.

In such retrogression, thinking would naturally not be taken back to precisely the point where it had been before. For the failure of a certain solution has at least the result that now one tries *"in another way."* While remaining in the framework of the old *Problemstellung,* one looks for another starting point. Or again, the original setting may itself be altered *in a definite direction,* because there is the newly added demand: From now on, that property of the unsatisfactory solution must be avoided which makes it incompatible with the given conditions. An example: The fully developed form of our radiation problem is naturally preceded by a stage in which the problem runs only as follows: Destroy the tumor with the aid of appropriate rays. The most obvious solution, which consists simply in sending a bundle of sufficiently strong rays through the body into the tumor, appears at once inadequate, since it would clearly have the result of destroying healthy tissue as well. In realization of this, *avoidance of the evil* has to be incorporated *as an additional demand* into the original form of the problem; only in this way does our form of the radiation problem arise (cure . . . without destruction of healthy tissue). One more example: In the pendulum problem, a watchman is often proposed who has the task of keeping the length constant by compensatory changes in the position of the weight. For the most part, Ss realize spontaneously that this procedure could not possibly be sufficiently precise, and that it would also incessantly interfere with the motion of the clock. Thus the problem: "compensation of the change in length of the pendulum," is enriched by the important addition: "automatically."

Such learning from errors plays as great a rôle in the solution-process as in everyday life.[13] While the simple realiza-

[13] Life is of course, among other things, a sum total of solution-processes which refer to innumerable problems, great and small. It goes without saying that of these only a small fraction emerge into consciousness. Character, so far as it is shaped by living, is of the type of a resultant solution.

tion, *that* something does not work, can lead only to some variation of the old method, the realization of *why* it does not work, the recognition of the *ground of the conflict,* results in a correspondingly definite *variation which corrects* the recognized defect.

13. FAMILY TREES OF TWO SOLUTIONS

As a supplement, there follow two family trees compiled from many combined protocols. That of the radiation problem is derived from experiments with German and American psychology students, of whom 16 were tested in individual experiments and 26 in group experiments. P_1, P_2 . . . denote successive phases of the process as they follow one another, not in chronological order, but ordered as to content. The Roman numerals and the large and small letters each denote coordinated phases. The numbers following the phases indicate how many times the phase concerned appeared in the protocols; those in parentheses how often with the E's occasional help, the rest how often spontaneously. If no number follows, the phase has been interpolated as expressing the functional value of solutions actually recorded.

The second family tree refers to the pendulum problem, and is derived from the records of six individual experiments and from 33 gathered in group experiments.

All the tentative solutions offered by the Ss are noted down in the family trees, including the foolish ones and those which are hardly comprehensible. Naturally, the records of group experiments are much poorer in mediating phases than those of individual experiments. The Ss in groups had only about ten minutes to work in; in the individual experiments, they had, where necessary, more than an hour. Moreover, the Ss in individual experiments were continually urged to think out loud, and did not themselves have to write down what occurred to them.

Family Tree of the Solution of the Radiation Problem

P_1 Destroy the tumor by means of rays without also destroying healthy tissue. 42
P_2 I No contact between rays and healthy tissue. 1
P_3 IA Redirect the rays over a path as free as possible from tissue. 2(2)
P_4 IA
a) . . . via the esophagus or intestines. 12
b) Swallow the radiating substance. 4
c) Send the rays over the shortest path. 7
d) . . . through an area less vitally important. 7[14]
e) Send the rays from above.[15]
P_3 IB Remove healthy tissue from path of rays by operating. 17
C Place a protecting wall between rays and healthy tissue. 1
P_4 IC a) Introduce a cannula. 8
b) Swallow a substance which is impermeable to the rays. 1
c) Coat the tumor with lead, leaving an opening for the rays. 1
d) Place an insulating plate on the surface of the body. 5
e) Surround the rays by harmless rays. 2[16]
P_2 II Desensitize the healthy tissue on the way. 8
(or the reverse: Render the tumor more sensitive. 4)
P_3 IIA . . . by chemical injection. 10
B . . . by previous immunizing application of radiation. 2
Or: Aren't pathological tissues perhaps more sensitive? 6(1)
Or: Apply rays to which pathological tissues are more sensitive. 9
P_2III Less intensity of radiation on the way. 6(6)

---

[14] These two suggestions came from another group experiment, in which by accident the ellipse in the accompanying sketch was made especially narrow. (See 3.)

[15] This S took literally the accompanying sketch of a cross-section.

[16] Proposals P_4 IC *b* and *c* clearly depend on a partially false setting of the problem. Not only the healthy part of the stomach, or the second half of the path, must be protected. Proposal *d* comes from a completely false conception of the task. Proposal *e* involves the application of a false model: the rays are treated as though they were poison which must not come in contact with the body.

P_3IIIA Have the rays less dense on the way. 2(2)

P_4IIIA
- a) By means of a focussing lens. 5(2)
- b) By crossing several weak bundles of rays. 2(2)
- c) By rotation of the body, or of the radiation apparatus, around the tumor. 1

B Decrease intensity on the way by opposing forces (for example, by other rays falling on the first at right angles). 12

C Weaken intensity on the way by interference. 3

D Use rays which work only at a definite distance? 3

E Give the rays full strength only when the tumor is reached. 1

P_2 IV Cure without rays. 1[17]

P_2 V Do not mind having the healthy tissue destroyed (it would be the lesser evil). 1

Family Tree of the Solution of the Pendulum Problem

P_1 Prevent irregular working of the clock which would follow alteration in the length of the pendulum caused by temperature-change. 39

P_2 I Constant temperature. 11

P_3 IA Protected place (prevent temperature-changes from reaching the clock). 3(4)

P_4 IA
- a) Clock in cellar. 1
- b) . . . Not on outer walls. 1
- c) . . . In a narrow space protected from draught. 1
- d) . . . In a vacuum. 4
- e) . . . In a hermetically sealed container. 5
- f) . . . In steam. 1
- g) . . . In a coating of ice. 1

P_3 IB Compensate the changes in temperature.

P_4 IB
- a) Someone to regulate the temperature by appropriate warming or cooling of the room. 1
- b) . . . By appropriately changing the location of the clock. 1
- c) Automatic introduction of compensating influences on the temperature (expansion into cold water). 1[18]

P_2 II Avoid the alterations of length caused by actual temperature-changes. 1

P_3 IIA Use insensitive material. 10(1)

P_4 IIA
- a) . . . Wood. 3
- b) . . . Platinum. 1
- c). . . A coating of appropriate varnish. 1

P_3 IIB Compensate the changes in length. 4(5)

P_4 IIB
- a) Someone to regulate (change location of) weight on the shaft. 10
- b) Combine simultaneous contractions (or expansions) in opposite directions. (Fig. 5A, h–p)
- c) . . . Simultaneous expansion and contraction. (Fig. 5A, q–s)
- d) Have the weight lifted by expansion of the shaft. (Fig. 5, t)
- e) The point of suspension can move up and down on the shaft (or chain). (Fig. 5A, u,v)
- f) The point of suspension shifts in the same direction as the center of gravity. 2[19]
- g) . . . In the opposite direction. . . .[20]
- h) Leaving leeway for change in length (as with railroad tracks). 1[21]

P_3 IIC Block the change in length.

P_4 IIC
- a) . . . By having it blocked by a spatial barrier. 1[22]
- b) . . . By keeping it restricted within a fixed framework. 2[23]

P_2III Prevent the effect which change in length would have on working of clock.

P_3IIIA A counter-weight above the point of suspension. 3

B Compensating electric or magnetic influences. 1

C Compensating change in the force of gravity by change in location of the clock. 1

D Compensating change of the clock-work. 1

E Proportional change of the whole system. 2

F A metal which retains its center of gravity despite displacement. 1[24]

G The weight to be very heavy, the shaft very light.[25]

H Change in the weight compensates change in length. 1[26]

14. DISCUSSION OF N. R. F. MAIER'S VIEWS

From N. R. F. Maier, the American psychologist, we have had in recent years several reports on admirable experiments

[17] This and the next proposal are attempts to escape from the original setting of the problem.

[18] This proposal is a particularly crass example of a solution which is 'single-track', i.e., leaves all other factors out of consideration. ("Operation successful, patient dead.")

[19] This is a crude error in thinking and an unfortunate attempt toward the proposal in *e*. For the displacement of the point of suspension in "absolute" space naturally has no effect on the length of the pendulum.

[20] This is the solution of a problem mistakenly confused with the original, namely, to keep the weight in a constant location.

[21] Not wholly clear.

[22] See Footnote 8, page 11.

[23] And the framework?

[24] Indeed a vain wish!

[25] Completely confused.

[26] But how?

in which the problem-solving of men and of rats was investigated. (26, 27, 28.) Maier belongs to the few psychologists who are now concerned with precisely the questions which are raised in the present chapter. On the basis of his experiments, Maier arrived at the following main statements about productive thinking:

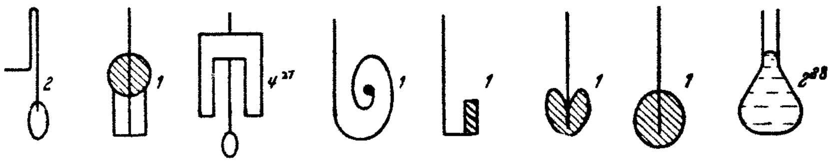


FIG. 5a, h-p

1. Thinking consists in *combining* parts of previous experiences *in new patterns*.

2. Such a new combination occurs always under the influence of a definite *'direction'*.

"The parts . . . must be combined in a certain manner, and a direction or way the problem is attacked seems to be a factor which determines the nature of the combination." (27), p. 143.

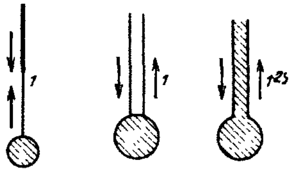


FIG. 5a, q-s

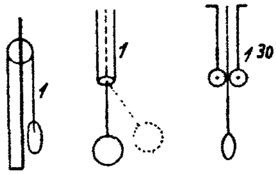


FIG. 5a, t-v

These theses derive principally from the following experimental procedure: Maier divides the solution of a problem in several parts, and gives these to his Ss as aids. He finds that these parts of the solution are of hardly any help, if the 'direction' is not also given.

Now when one closely examines what Maier calls 'direction', it becomes evident that *'direction' is nothing but the earliest phase of the solution, i.e., the reformulation of the problem as it initiates the solution-process concerned. The other 'aids' contain later (subordinate) properties of the solution.* The reader may convince himself that this is the correct interpretation of Maier's views by considering the following example: "There are several angles or directions, that one may take from the same starting point. . . . If two doctors wished to solve the yellow-fever problem they might see the difficulty altogether differently. One might believe that the cure depended upon making the human immune to the germ [direction 1], the other might think that the germ must be kept from the hu-

[27] This solution was familiar in all four instances. In the case of the two preceding solutions, the Ss concerned remarked that "something of this sort" was known to them.

[28] Containing mercury (or instead of this: a column of mercury lifts a pendulum-weight from without).

[29] Expansion and contraction are here supposed to "hinder each other".

[30] Here the pendulum swings perpendicular to the plane of the drawings. This proposal of course makes the mistake that the "wheels" could not move along the pendulum-shaft, which is fixed nowhere else.

man [direction 2]. The first might experiment with serums, the other might seek to learn what carried the germ and hence seek a way to remove the carrier." (27), p. 137.

From the comparison of Maier's concepts with those of the present investigation, it follows that:

1. There is no need for a special concept of a 'direction' which combines elements. 'Direction' is entirely of the type of a *problem,* or, more exactly, of the reformulation of a problem and of a mediating phase in the solution-process. Let the yellow-fever problem begin or end with one of the two 'directions'; then the problem- or solution-character of the 'direction' involved becomes clearly evident. In short, what Maier calls 'direction' is an 'organizing principle' in no other sense than is what he calls 'problem'.

2. There exists as little fundamental difference between 'direction' and 'the elements to be combined' as between 'direction' and 'problem'.[31] For these elements combine with one another with only apparent simultaneity. In reality, they usually follow upon one another in a sequence in which the each element possesses problem-character (thus, 'direction'-character) with respect to the following, and solution-character with respect to the preceding elements. In other words, the combination is carried out in the form of a succession in which more particular realization-aspects or means-aspects of the solution develop from functional values or goal-aspects. Incidentally, the same follows from Maier's rat experiments.

I have spent so much time in the reinterpretation of Maier's concepts for the reason that material of a given type should not for long be expressed by different concepts.[32]

[31] That these elements need not under all circumstances be "parts of *past* experiences" is shown at least by all those solutions into which newly offered objects enter according to their suitable properties.

[32] For the same reason, Claparède's concept of *hypothesis* (2) appears to me superfluous. For Claparède's "hypothèse" is nothing but a solution or a preliminary, provisional phase of a solution.

Chapter II

The Solution of Practical Problems (II)

1. THE ASSOCIATION THEORY OF PROBLEM-SOLVING

The main contribution of the preceding discussion was this: the solution of a new problem typically takes place in successive phases which (with the exception of the first phase) have, in retrospect, the character of a solution and (with the exception of the last phase), in prospect, that of a problem. The problem we are considering takes therefore the following form: *How does a solution-phase arise out of the immediately preceding problem-phase?* It is probably clear that the solution cannot take place reproductively by virtue of mere *"associations"* among the contents of the various phases. The explanation by association, moreover, becomes no more plausible if one adds the thesis that not these identical contents, but only similar ones, have been previously associated, and that this suffices. Let it be kept in mind that the classical concept of association has no reference to any such "material" relations[33] between associated contents as "cause of . . .", or "solution of . . .", but solely to temporal and spatial contiguity or similarity. But between a problem or its several parts and a solution there is no more spatial and temporal contiguity and no more similarity than between that problem and innumerable other contents, for instance, similar problems, or any circumstances accompanying earlier solution-processes, and so on. It follows that, according to the association theory, completely nonsensical errors should frequently occur in the solution of thinking problems. For example, the S should, in good faith, actually name as a solution some event which happened to take place on the occasion of an earlier solution, or he should mention a similar problem, *as though it were a solution.* From the fact that such errors do not occur, O. Selz (37) has already drawn the necessary conclusions against the classical association theory of thinking, against the theory of "diffuse" association and reproduction.

2. FINDING A SOLUTION THROUGH "RESONANCE-EFFECT"

A reproduction theory of the type of Selz's, in which room is expressly left for such relations as, e.g., "part of . . .", "next to . . .", "origin of . . .", "solution of . . .", deserves much more serious consideration. Without doubt, through experience, events can acquire the attribute, "solution of problem A", or at least, "leading to effect *a*". Now if at some time something with the attribute, "solution of problem A" or "leading to effect *a*" is sought, the solution can be found *by virtue of the correspondence between the attribute desired and that which inheres in what is sought.* (Compare Selz's concept of "determined means-abstraction".) In contradiction to the classical reproduction theory, the solution is here found by the "anticipation"[34] or the "signalling" of its specific solution- or means-properties in each case. (It need not even have occurred—or have become familiar—as a solution. It must merely have been experienced as

[33] ["Sachbeziehungen"]

[34] In seeking, certain attributes of what is sought are "anticipated" or, as we should prefer to say instead, "demanded", "signalled".

"leading to" an effect similar to that which is now, for the first time, the goal. An example: For certain experiments in the psychology of perception, someone needs yellow illumination. There is no color-filter available. What to do? It occurs to him how, the other day, a blue folder reflected the light of a lamp as blue-tinted, "led to" a coloring. Aha! The reflection from yellow paper. . .)[35]

In agreement with Selz's theory of "schematic anticipation," we can write the general formula of such problem-solving as follows: The problem is: $?Rb$[36]; aRb exists in the thinker's experience; by reason of the partial correspondence with $?Rb$, aRb and therefore a are aroused. Thus this finding of the solution takes place ultimately through a kind of "excitation by equality" (Selz) or, better, of *resonance*. (See details in Chapter VI.)

Let us now ask ourselves whether the genesis of the solution of a new problem can be quite generally explained by the resonance-effect of an appropriate signal. That the problem and the solution are new, would be no argument against this. We have already seen that the solution need not have occurred before as 'solution'. Moreover, the several solution-phases might always represent familiar solutions, while their *combination* would be new. In our examples, the radiation problem and its solution are new, but the first solution-phase: "intensity small on the way, great in the tumor", represents perhaps, in its formal characteristics, a familiar solution of the more general problem: "Find a means to achieve at a certain place an effect which is to be avoided on the way." Do we not often decrease the intensity of some agent, when a diminution of the effect concerned is desired? And the "diffusion of the agent on the way", in its turn, would again be a familiar solution of the problem: "a means to achieve less intensity".

Obviously, the theory of the "solutions by the resonance-effect of signals" deserves detailed examination. In point of fact, there are solution-phases everywhere which come about in such a way. At the end of a solution-process, for example, some appropriate object is often still to be found whose introduction into an already discovered procedure completes the solution. Thus in the radiation problem, one may be after a "free path to the stomach", and find the esophagus. Or a chimpanzee may look for something "long and movable" (with which to fish for the banana) and find a branch or a piece of wire or the like. Such parts of the solutions, no matter whether they are found in the perceptual field (as in the last example) or in memory (the field of traces), really originate by reason of the correspondence between the attribute desired and that inherent in the object. ($?Rb$ would here mean that a thing is sought which has the attribute b. R would therefore symbolize the relation of an object to one of its attributes.)

In Chapter VI, this origin of a solution will be discussed in detail. At present we are interested only in this: Can the origin of even the earlier solution-phases, in which no definite real object but only the "procedure" is yet to be found, also and always be interpreted after the same pattern? It seems advisable to phrase the question thus: *Must* the origin of *all* phases be thus interpreted?

Of what aspects of a solution, then, does the theory just described really make use? If the solution meant nothing but "something which leads to the goal"—so

[35] This "leading to . . ." is very often given as immediate perceptual experience (cf. Chapter V).

[36] R signifies some relation, e.g., that of "leading to . . ."; therefore aRb signifies a in the relation R to b.

that merely "if S, then G, and if not S, then also not G"—then such a theory of problem-solving would still be applicable. For this is all the theory requires. *Aside from this, under this theory, the contents of solution and of goal could be in any relation whatsoever.*

Suppose that we could interchange the solutions of two different problems (e.g., that of the radiation problem and that of the pendulum problem), in such a way that S_1 henceforth leads to G_2, and S_2 to G_1. A living creature who had never had experiences to the contrary would learn this; that is to say, for this creature, S_1 would acquire the attribute of being *a way to* G_2, and so on. Suppose now that the finding of a solution required nothing but an anticipation of this leading-to character in question. When one day the problem arose to find a way to G_2, this creature would be able to find S_1 just as readily and in the same way as another creature with normal past experience would find S_2.

In other words, a theory according to which a solution is found by reason of the resonance-effect of a suitable signal would be equally applicable to any conceivable number of combinations of problems with solutions. The question arises: does the *actual* relation of a solution to its problem offer no other heuristic possibilities?

3. HEURISTIC METHODS OF THINKING: ANALYSIS OF THE SITUATION AS ANALYSIS OF CONFLICT

Let us investigate how a solution is actually related to its problem. We find that a *solution always consists in a variation of some crucial element of the situation.* Thus, in the solutions of the radiation problem, we find changed: either the spatial relations of rays, tumor and healthy tissue; or the intensity or the density of the rays; or the sensitivity of the tissue. And in the first case, the position either of the rays, or of the healthy tissue, or of the tumor can be varied. With this, the primary *conflict-elements* of the radiation problem are probably exhausted.

Thus every solution takes place, so to speak, on the concrete, specific substratum of its problem situation. For this obvious reason, quite apart from other factors, every solution of the radiation problem, *qua* solution of precisely this problem, is differentiated from every solution of, say, the pendulum problem. Solutions of the latter problem have to do with temperature and with the structure of the pendulum, not with tissues, rays, and the like. This is as important as it seems to be banal. For it follows from this that, in seeking a solution, one must bring the given problem-situation as clearly as possible into focus. He who merely searches his memory for a "solution of that such-and-such problem" may remain just as blind to the inner nature of the problem-situation before him as a person who, instead of thinking himself, refers the problem to an intelligent acquaintance or to an encyclopedia. Truly, these methods are not to be despised; for they have a certain heuristic value, and one can arrive at solutions in that fashion. But such problem-solving has little to do with thinking.

With the radiation problem as well as with the pendulum problem, it happened that Ss who were already dimly familiar with the solution asked themselves, "I must have heard of this before —now, how was it?" Every reader will already have experienced that such seeking-to-remember will often (say, in mathematical problems) disturb new thinking.

We can therefore say that *"insistent"*

analyses of the situation, especially the endeavor to vary appropriate elements meaningfully sub specie of the goal, must belong to the essential nature of a solution through thinking. We may call such relatively general procedures, *"heuristic methods of thinking"*.[37]

The inquiry after elements which should be varied in a suitable fashion is identical with the question, *"Just why doesn't it work?"* or, *"What is the ground of the trouble (the conflict)?"* For, each component of the situation whose variation means a solution is in its original form a "ground of conflict" (e.g., the great intensity of radiation en route, or the spatial coincidence of rays and healthy tissue). To each solution corresponds a ground of conflict present in the situation. Analysis of the situation is therefore primarily *analysis of conflict.* In connection with the radiation problem, the S will ask himself, "Just why are healthy tissues destroyed as well? What elements of the situation are responsible for it?" In other words, he does not merely inquire after a way to avoid the accompanying destruction of healthy tissue, but *seeks to penetrate more deeply into the nature, into the grounds of the conflict.*

The statement of the radiation problem contains, moreover, virtually all the causal relations which are necessary for recognition of the immediate grounds of conflict and for devising the corresponding solution-phase. In setting the problem, it is said, ". . . rays, which at sufficient intensity destroy organic tissue . . ."; the dependence of the effect on the agent's intensity is here asserted. There is further given the relation of the effect to the sensitivity of the tissue, for it is said, ". . . rays, which . . . destroy organic tissue . . .". This involves the category of an effect's dependence on the nature of the affected object, and so on.

4. ANALYSIS OF SITUATION AS ANALYSIS OF MATERIAL

Of course, analysis of the situation has not only the meaning of analysis of conflict. In general, the problem situation contains also, more or less completely and explicitly, all sorts of *material* for the various solutions. Besides elements which, in the solution, undergo elimination or alteration (so-called conflict elements), there are also those which are actually used by the solution (*material elements*). What we earlier called "suggestion from below" (p. 11) is founded on the relatively spontaneous action of the latter. While the conflict elements answer the question, "What is wrong here? What must I change?", the material answers the question, *"What can I use?" Analysis of situation appears therefore in two forms: as analysis of conflict and as analysis of material.*

Let us illustrate by several observations—which supplement the examples in Chapter I, 11—how certain parts and features of the problem-situation are taken up as "material" by the solution-process. With the problem: to find "an absolutely unfailing signal to send down a river, i.e., one which cannot catch or be interrupted on the way", a S was struck by the fact that water slips through irresistably everywhere, and thus found the solution: "coloring of the water." Again, in the preparation of one of the group experiments on the radiation problem, the sketch of the ellipse turned out rather narrow. This had the

[37] It is true that these heuristic methods, in their turn, could be aroused by way of mere resonance, by the very general anticipation of: "an appropriate method of solution". It will be shown in Chapter IV, however, that over and beyond this there is a more 'insightful' way to these methods.

result that, out of 43 who took part in this group experiment, 7 proposed to send the rays by the shortest way (i.e., in the direction of the short axis of the ellipse, instead of the long axis). I encountered this proposal nowhere else. Obviously the short axis had suggested itself as a more suitable path.

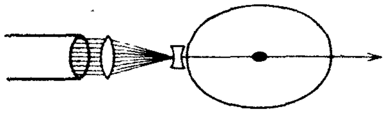


FIG. 6

In connection with the radiation problem, I once gave a group of 5 Ss a sketch which showed, in a position of secondary importance, a lens and a condensation of rays (see Fig. 6)—as though the radiating cylinder were by chance too large, and a convex lens were therefore necessary to reduce the bundle of rays to a smaller diameter.[38] For comparison, I gave to 5 other Ss the usual sketch. The results were: 3 of the 5 Ss in the "lens" *group incorporated the lens into the* solution, while none of the comparison group happened on the lens-solution. According to my other experiences, indeed, not 5 but 75 Ss would be necessary to produce 3 lens-solutions without the help of a lens in the drawing. The lens and the condensation of rays in the drawing suggested, therefore, their appropriate use.

An analogous effect, namely, the inclusion of explicit components of the situation in the solution, was obtained from an experiment in which 12 Ss received another sketch. It contained, in addition to the destroying bundle of rays, yet another bundle, of X-rays, falling at right angles to the first—allegedly "to determine the position and the condition of the tumor". 12 other Ss, to whom the usual sketch was given, constituted the comparison group. 6 of the 12 Ss in the "X-ray" group hit upon solutions which made use of 'other' rays (in part, explicitly of X-rays), "to abolish the harmful effect of the rays" and the like; only 3 of the 12 comparison Ss proceeded in such a way.[39]

We should like to illustrate the adoption of, and application given to, attributes of the situation with yet another and completely different problem-material. The problem—which N. R. F. Maier has already used in an actually realized situation—runs as follows: In a large room, two ropes hang from the ceiling at a considerable distance from one another. One has a small ring on its free end, the other a small hook. A S receives *the* task of *fastening the* two ropes together, but this is not possible directly. For the problem is just this: To begin with, how is he to get both ropes in his hands at once? The ropes hang inconveniently far apart.

Typical solutions of this simple problem follow:

1. He goes with one rope to the middle, fastens it here on a chair or the like, and then gets the other one.
2. He takes a long stick and fishes for the other rope.
3. He fastens a cord to one rope, and can thus lay hold on the other without releasing the first.

[38] See page 3, footnote 4, in reference to "lens".

[39] By the way, of the 12 Ss of the X-ray group (and also of 4 other individual Ss to whom the radiation problem was given with the X-ray sketch), not one was stimulated to the solution: "crossing of weak bundles of rays" (to which end the crossing X-rays had been added). This must at first appear surprising, since we only just saw how thoroughly such a model is usually "ransacked" for means to a solution. The reason is probably that the difference of their given function from that of the main rays is too firmly impressed on the X-rays. (Compare the concept of "shift of function within a system" in Chapters VII and VIII.)

4. He sets one rope (or both) to swinging, and catches it as it swings towards him. Or, he swings himself on one rope over to the other.

5. He climbs a ladder, since a rope-end comes nearest to the other rope in the vicinity of the ceiling.

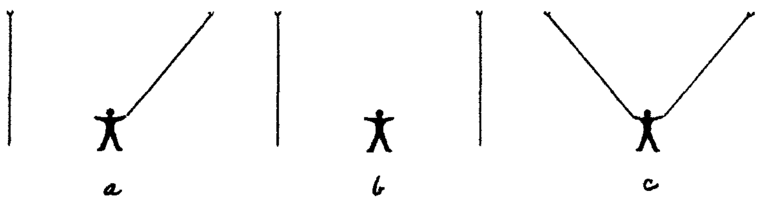


FIG. 7

If various individuals are given the foregoing problem, some will start from such a visualized model as Fig. 7a, others from a model such as 7b, or, again, 7c. We now distributed these three models among three groups of Ss as "starting models," because we wished to see whether such details of the model would exercise a noticeable influence on the direction of solution. The result was that solution 1, the "anchoring" method, occurred to 12 out of 18 Ss with model 7a, i.e., in 66% of their protocols. Over against this, the "anchoring" method appeared with only 1 out of 7 Ss when model 7b was given, and with only 4 out of 13 Ss who had model 7c, thus in only 25% of these two groups taken together. Clearly, model 1 suggested most directly the anchoring of the rope ("If it remained here, then . . .").

In the papers of N. R. F. Maier already mentioned (26, 27, 28), one finds fine experimental data on the effect of *aids*, both of those which supply suitable material to the "directions" already present and of those which suggest new directions.

5. ANALYSIS OF GOAL

Next to analysis of the situation in its two forms, as analysis of conflict and of material, a genuine thinking process is characterized by the *analysis of the goal, of what is demanded,* by the question: *"What do I really want?"* and perhaps by the supplementary question: *"What can I dispense with?"* In this way it can become clear, for instance in the radiation problem, that the rays' progress in one bundle, as represented in the unsuitable starting-model, is not demanded at all and is hence dispensable. In the pendulum problem, the normal shape of the pendulum can analogously be recognized as "dispensable". By means of the question: "What is *really* demanded?", thinking is able to free itself from hindering "fixations".

A similar rôle is played by the intentional generalization of the problem, of the goal, by the question: "What does one do, in general, when one . . . ?" In the radiation problem, when a S "no longer saw the forest for trees", I often recommended this heuristic method of generalization with the words, "Well, what does one quite generally do, when an agent is intended to achieve an effect at a certain place while this effect should be avoided on the way to this place?" To be sure, the S then very frequently said, "But that I have been trying to do the whole time". Nevertheless, it helped him, so to speak, to "loosen" his ideas, to remove fixations.

Later on, in discussing solutions of mathematical problems (see Chapter III), we shall become acquainted with still other forms of analysis of the situation and above all of the goal. In the meantime, we may summarize the results which have been achieved in this chapter so far. *In a genuine thinking process, certain heuristic methods play a decisive rôle in mediating the genesis of successive solution-phases.*

These heuristic methods are not cited in the genealogies given. They represent no solution-phases in the sense of those cited; they are not properties of the solutions sought, but rather "ways" to find them. They ask: "How shall I find the solution", not: "How shall I attain the goal?"[40]

Intelligent heuristic methods can be observed even in the most primitive animal experiment. Thorndike, in his famous experiments on cats (see page 10, or 39), could in this way establish that the "tendency to pay attention to what it is doing" (instead of blindly struggling) increased in the course of the experiment. And in experiments with the same problem situation, Adams (1, p. 92) found that: "Usually there was little time spent in activity as compared with the amount of time spent in looking over the situation."

6. DISPOSABILITY (LOOSENESS) OF ELEMENTS IN A SITUATION

What direction a solution-process takes at every moment depends on the psychological relief-map of the problem situation, on the "disposability", the "looseness" of the elements concerned. Many Ss conceive of the radiation problem, at least at first, as if unconditionally and solely the path of the bundle of rays had to be suitably varied. The other crucial elements (such as the intensity of radiation, the inner condition of the tissue) remain taken for granted, "firm", "not thematic" [i.e., not belonging to the "theme", the center of operations in the situation, the "figure"].

The following experiments show on what insignificant nuances of the *Problemstellung* the direction of a solution-phase may depend. Two groups of Ss were given the radiation problem in the same words and with the same sketch, except that just two sentences, which were to explain the inadequacy of the direct solution, were differently formulated for the two groups.

Group I received the sentences: "The rays would thus destroy healthy tissue, too. How could one prevent the rays from injuring the healthy tissue?" In place of this, Group II received the sentences: "Healthy tissue, too, would thus be destroyed. How could one protect the healthy tissue from being injured by the rays?" In short, the identical (!) problem-content is expressed once in the active case, and then in the passive. In the first case, the emphasis lies on the rays, in the second on the healthy tissues. To determine whether the direction of solution, the choice of a genealogical line, is noticeably influenced by such a difference in emphasis, I counted in both groups the protocols in which intensity of radiation was in some way made the point of attack. All lines issuing from P_2III—see page 14—would belong here, as well as the question: "Are diseased tissues possibly more sensitive, so that one could use weaker rays?"—in short, all "solutions by intensity". The frequency of such a procedure may doubtless be taken as indicating a special "looseness" of the rays.

Actually, the result was that the "ray group" dealt with the intensity of radia-

[40] A solution is of the type of a way to the goal; a heuristic method, on the other hand, is of the type of a way to solutions.

tion in 10 out of 23 protocols (43%), the "tissue group" in only 3 out of 21 protocols (14%). Aside from this, in the first group the intensity of radiation was much more dominant where it did appear.

For almost all Ss, the rays' "being in a single bundle" (*one* bundle projected from *one* tube) was to such a degree taken-for-granted, fixed condition of the situation that, for this reason alone, the solution: "concentration of several weak bundles of rays in the tumor," could hardly arise at all.[41] Had I recognized the importance of this early enough, I should rather have omitted the sketch in the main experiments. It proved disturbing because it caused fixation. As it was, I puzzled long over the fact that the solution-phase: "concentration of rays," or its three subordinate final solutions P_4III A a-c (see page 15), so seldom appeared. For example, in 26 records of group experiments, only two spontaneous statements were made which fell within these four phases; moreover, both were made by the same S. The other group experiments with the radiation problem yielded no better percentage. Even though the Ss in all group experiments had at most only 5 to 10 minutes' time, yet the small number of "solutions by concentration" appears surprising. Eventually my suspicion fell firstly on the sketch and secondly on the ray-nature of the agent. (That intensities of radiation may be added up might not be familiar enough to some Ss. Phenomenally, rays do not as a rule fall easily into separable parts which can be added up.) Several group experiments were set up to test this suspicion.

1. 11 Ss received the radiation problem *with the sketch,* 11 other Ss *without the sketch.* (The Ss were students of about 16 and 17 years of age, in a secondary school.) Result: with sketch, 9% solutions; without sketch, 36% solutions.

2. In two other group experiments *without the sketch,* 28 Ss in all received the original statement of the problem (given in abbreviated form on page 1), while 30 Ss received the following variation (which substitutes *"particles"* for "rays"):

"Assume that an individual has a tumor within his body, let us say in the stomach. Assume further the existence of a kind of tiny missiles (perhaps of the size of atoms) which have the power to penetrate and therewith to destroy organic tissue, when they strike it in sufficient number per unit of space and time. With the help of such tiny missiles, how could one free the patient of his stomach tumor?"

Result: with the ray model, 18% solutions; with the particle model, 37% solutions. (These Ss were partly university students, partly students of 15 or 16 in a secondary school.) Thus the suspicion was confirmed in both respects.

A conflict-element can obviously exhibit very different degrees of fixedness. There is a degree of fixedness which is sufficient to divert the tendency towards change from this element to places more yielding, but which would yield to a relatively small pressure toward loosening up ("Now I'll try it differently."), and to a systematic probing for "other places for attack." On the other hand, there is a degree of fixedness which is a match for practically all pressures; in

[41] Here statements often occurred such as, "Well, I thought of course I had to do it with *one* ray apparatus, as in the sketch." Analogously, there were Ss for whom, in the pendulum problem, the common form of a pendulum was untouchable, so that solutions of the type P_4II B b-e (see page 15) had no great chance with them. (Occasionally, this followed from the fact that they worked with a cord instead of a shaft, for the pendulum.)

this case we speak of "fixation." A drastic example is furnished by that familiar match problem in which four equilateral triangles are to be constructed out of six matches. The solution is a tetrahedron, an object in three dimensions. We tested five Ss individually and perhaps 40 in group experiments. Certainly for all Ss the original situation was "the same," at least in so far as all of them worked at

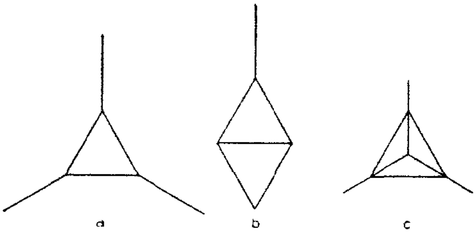


FIG. 8

first with a plane model, as if the problem had been, ". . . *lay* four equilateral triangles. . . ."[42]

But the firmness of attachment to the plane differed enormously with different individuals. For some Ss, the figure jumped with a sudden jerk into the third dimension, if certain favorable constellations occurred such as appear not infrequently in the search for a solution—an important function of "intelligent fumbling." This is much facilitated by a sort of "satiation"[43] with regard to the plane after protracted, fruitless search therein. Such fortunate "suggestions from below" are supplied, for example, by the three plane constellations of Fig. 8, a-c. Yet most Ss cannot be helped even by such prodding; they are and remain irreparably "fixed" on the plane.

It should be noted that the "fixedness-relief" of a problem situation does not depend on arbitrary distribution of the Ss attention. On the contrary, his attention is for the most part determined by the given "fixedness-relief."

7. CONDITIONS OF THE DISPOSABILITY OF ELEMENTS IN THE SITUATION

In the following, we shall investigate what factors—*aside from knowledge and habit*—determine the "looseness" of an element (specifically, of a conflict-element), its "becoming thematic." One of these factors is shown up nicely in the following primitive problem situation. A monkey wishes to go through a narrow door with a stick. He does not get through. What to do? The primary conflict-element is the collision, the crossing of stick and doorframe. This relation can be recognized as the ground of conflict with especial ease, since a stage preceded in which it was lacking and unhindered movement occurred at the same time. Without crossing there was progress, with crossing there is no longer progress, i.e., *both sides of the causal relationship are realised successively in one and the same situation.* Therewith, the conflict element acquires the immediate perceptual character of an 'obstacle.' We can give this *first* condition the following general form: a conflict element may be recognized with particular ease on the background of its opposite.

I borrow a further example for this from the "door" problem: a door is to be constructed so as to open toward both sides. How can this be attained? (Fig. 9 shows a cross-section through adjoining pieces of door and wall, and the hinge.) The fact that "the door does not open toward the right" has two grounds of conflict: "The right corner of the door cannot turn further on its way" and "the left corner of the door does not leave the

[42] This is clearly a result of the fact that all other familiar match problems proceeded on a plane, which must naturally lead to a blending of the character: "match problem" with the character of "laying." Aside from this, matches are by nature not fitted for building in three dimensions.

[43] Cf. A. Karsten (17). Karsten observed, with growing "satiation", an increasing tendency to spontaneous variations of the activity involved.

wall." Of these the first is clearly more striking, more thematic than the second, because it implies the sudden interruption of a movement now in progress, while the opposite of the second has not yet appeared in any way. The greater actuality of the first-named conflict may partly have accounted for the fact that, with six out of seven Ss, ideas for a solution which were centered about removal of the first conflict element—specifically, the bevelling or the thinning of door and wall—preceded ideas centered about removal of the second conflict element—for example, a rubber hinge, or two hinges of which the left snaps out when the right snaps in.[44]

A *second* factor for a conflict element's becoming thematic is illustrated again in the "stick-door" task (p. 26). As solutions of the problem-phase: how the disturb-

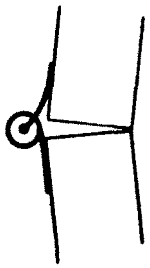


FIG. 9

ing crossing of doorframe and stick is to be avoided, in principle at least the following four variations are possible:

1. Rotation of the stick into a position sufficiently vertical or normal, i.e., perpendicular to the plane of the door,
2. Shortening of the stick, with oblique position unaltered,
3. Widening of the door with oblique position and length of stick unaltered,
4. Breaking through the doorframe (or breaking the stick) at the point of crossing, with oblique position, length of stick, and width of door unaltered.

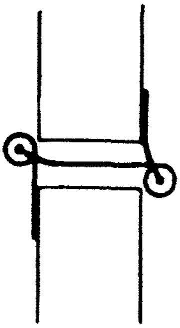


FIG. 10

To these four solutions correspond just as many conflict elements in the original situation. Now, by far the most evident of these is the spatial position of the stick. For it is the most fortuitous, the most variable, and nothing in the situation requires just this position. In other words, a conflict element is the more easily recognized, the more fortuitous it is.

A *third* factor: In the "door" problem the thickness and angularity of the adjacent parts of wall and door constitute an especially conspicuous conflict element, because the corners collide in an attempt to open the door, i.e., the corners suffer the conflict directly. One can say in general that a conflict element is the more easily recognized, the more, and the more directly, it participates in the conflict-event.

Examples from other problems: The difficulty of the "solutions by dispersion" in the radiation problem depend principally on the fact that here the decisive conflict element, the rays' being-sent-from-a-single-tube, is so far removed from the direct conflict-event: the destruction of healthy tissue, and from its direct cause: the great intensity of radiation in the healthy body. Many Ss per-

[44] The real solution lies in the same direction (see Fig. 10): the left hinge serves the rotation to the left and the right hinge is taken along; the right hinge serves the rotation to the right, and the left remains functionless.

ceived this great intensity as the real trouble, but believed that this was indispensable for the goal: destruction of the tumor. ("The rays must surely be strong and reach the tumor from without. Nothing can be done about that.") They did not see that the goal (destruction of the tumor) demands the great intensity en route only with the tacit assumption of merely *one* source of radiation. The difficulty is that the unity of the source constitutes a relatively remote conflict element.

Analogous facts are again and again reported from animal experiments, for instance by Köhler (20, pp. 245, 246): "A stout cord is fastened to a stick with which the animal could reach the objective; on the free end of the cord is a metal ring . . . which is slipped over a nail. . ." "Grande, Chica, Rana and Tercera first pull at the stick and endeavour persistently to solve the connexion 'rope-stick' . . ." (instead of attacking in the ring-nail region). This is certainly due in part to the fact that the crucial element "ring over nail" is, even in a purely spatial sense, so far removed from the immediate region of conflict ("the stick moves no further"). Yerkes, among others, makes similar reports.

The factors which are here effective are naturally not exhausted with those enumerated. The disposability of elements which may be considered as "material" is to be dealt with specifically in Chapter VII. It hardly needs special mention that, along with such factors, *habit* often decides what elements of a situation are preferably laid hold on. Earlier successes leave behind traces which co-determine the future "localization of the difficulty" in kindred situations. But it would not do simply to regard disposability, or its opposite: fixation, *only* as the effect of customary preference, as does, for instance, Maier, in an otherwise excellent investigation (29).

8. ON PAR-FORCE-SOLUTIONS

In this connection, let me describe a strange form of problem-solving, which I was at first inclined to consider as a mere curiosity, until I recognized that it represented a practically very important form of solution-genesis. With many tasks, one hits on the solution simply because, at a somewhat violent attempt to reach the goal directly, the physical situation gives way in the direction of the solution, and a directly affected conflict element disappears. Examples: Suppose that, in the door problem, one attempts forcibly enough to carry out the desired rotation on the original model. This need not be done in the real situation, but might take place more successfully in imagination, because here the effect can be exaggerated. Under certain circumstances the corners must then give way. That is, the solution by bevelling the corners may simply be read off from the known physical behavior of the model under the pressure of a solution by force. In just this way (with the same solution by force), the hinge may break or bend, which often leads directly to the solutions: "rubber hinge" or "hinge which snaps out at the proper moment."

Analogously with the related stick-door problem, if the stick is run through the door with sufficient strength, it gives way in the direction of certain (primitive) solutions: the ends break off.

A third example: with the pendulum problem, a S once imagined that the pendulum expanded while its end-points were kept firmly in place—with the result that the pendulum shaft bent out. Even if this "evasion of Nature" was not directly applicable, nevertheless it could at least suggest forms of solution which make use of expansion in other (opposite) directions.

This means that, in solving problems, one may very often be fruitfully guided by natural reactions to forcible demands in the direction of the problem's goal (method of solution-by-force).

Incidentally, the par-force-solutions are only exaggerations of a factor with which we became acquainted above (7, p. 27) as the third factor. Just as a conflict-event singles out certain conflict elements through the fact that it plays directly on them, concerns them directly, it can also affect these conflict elements so strongly that it "overruns" them and thus already causes a part of their elimination.

9. RESTRUCTURING OF THOUGHT-MATERIAL

Every solution consists in some alteration of the given situation. But not only this or that in the situation is changed, i.e., not only such alterations take place as one would have to mention in a simple commonsense description; over and beyond this the *psychological structure* of the situation as a *whole* or of certain significant parts is changed. Such alterations are called *restructurations*.

In the course of a solution-process, the "emphasis-relief" of the situation, its "figure-ground" relief, for example, is restructured in this way. Parts and elements of the situation which, psychologically speaking, were either hardly in existence or remained in the background —unthematic—suddenly emerge, become the main point, the theme, the "figure." Of course, the reverse also happens (see 6, 7 above).

Aside from the emphasis, the material properties or "functions" of parts are changed as well. The newly emerging parts of the situation owe their prominence to certain relatively general functions: this one becomes an 'obstacle,' a 'point of attack' (conflict element), that other a 'tool,' etc. At the same time, the more specific functions also change. For example, the esophagus becomes a 'passage for rays' or a triangle of matches becomes the 'base of a tetrahedron.'

Especially radical restructurations tend to take place in the *nexus or context of the whole*. Parts of the situation which were formerly separated as parts of different wholes, or had no specific relation although parts of the same whole, may be united in *one* new whole. For example, in certain solutions of the pendulum problem, the place of suspension may thus enter into some relation with the pendulum's length, while previously no psychological connection existed between these two parts. The next chapter will furnish more convincing examples of alteration of belongingness and grouping.

It has often been pointed out that such restructurations play an important rôle in thinking, in problem-solving. The decisive points in thought-processes, the moments of sudden comprehension, of the "Aha!," of the new, are always at the same time moments in which such a sudden restructuring of the thought-material takes place, in which something "tips over." In the third part of the present investigation, entitled "Fixedness of the Thought-Material," we shall treat in detail the opposition which an old structuring of thought-material may offer to a new and more adequate structuring of the same material. There we shall also indicate more exactly what kinds of "structuration" are essential in various types of solution-processes, e.g., in finding material objects needed in a solution (Chapter VII) or in finding a mathematical proof (Chapter VIII). It seems very probable that the greatest differences between individuals as to so-called "intelligence", "capacity for thinking", are based on differences in the facility of such restructurings. Köhler (22) and Lewin (24) have pointed to the possibility that the ease with which thought-material can be restructured depends upon certain properties of the "neural milieu" in which the processes involved take place, or indeed upon the "person's mental material" in general. At the end of Chapter VIII, we shall become acquainted with still another hypothesis.

But while, with undoubted justification, great stress was laid on the significance of restructuration or reorganization in thinking, another side of the problem was almost completely lost from view. *In what way* do these restructurations, and with them the solution, arise? *That* they take place, and that they happen more easily with one person, less easily with another, still in no way discloses *why* they take place, i.e., whence the directive "forces" derive which lead a thought-material from the old over into the new, adequate structuration.

In themselves, restructurations of a field can have the most varied origins. The psychology of perception is acquainted with a great number of reversible figures, which, simply through the fact that one lets them persist long enough in one of the possible structurings, tend of their own accord to "tip over" into the other, opposed structuration. Cf. the spontaneous interchange between front and rear in the perspective drawing of a cube, or a flight of stairs, etc. Here "satiation" is probably the origin of the restructuration.—Or, if one experiences in succession, or side by side, a series of partially identical complexes —whose objectively common component has not been made unrecognizable by camouflage—it happens under suitable conditions that the common part, the common aspect, becomes predominant, while in each individual case, taken alone, quite different aspects stand out phenomenally. (See Chapter V, 2) Thus the restructuration here occurs through a sort of "precipitation of common elements" (Wm. James' "abstraction by varying concomitants", 16).—Or, one searches for something with definite properties, e.g., something long and solid, and now everything which objectively possesses such properties is "centered" accordingly. Here the restructuration takes place through the resonance-effect of an appropriate signal (see Chap. II, 2 and Chap. VI)—Furthermore, a restructuration may be caused by intentional alteration of given perceptural structures.

In its first two parts, the present investigation has the task of revealing the *causal (directive)* factors in the genesis of a solution. An exhaustive description of all occurring restructurations was not intended, the less since in this respect essential work had already been accomplished. On the other hand, the gestalt theory of thinking had not yet approached the causal problems involved. The theory of "closure" or *"prägnanz"* is much too general to be of any great use here. Of course, the solution-process continues until "the gap is closed," "the organization is complete," "the disturbance is removed," "equilibrium or release of tension is attained." And this is undoubtedly relevant from the dynamical or, more correctly, the energy point of view. But of what kinds of events this "tendency toward equilibrium" or toward *"prägnanz"* can make use, this is the problem which must now be investigated in the light of gestalt psychology.

Chapter III

On Solution-Processes with Mathematical Problems

1. THE "13" PROBLEM

THE FINAL solution of a mathematical problem, specifically of a problem in which a proof is demanded, has the form: "something, not to be proved here, from which the proposition follows". With such mathematical problems as well, the solution is typically reached not in one step, but in several successive steps. Here again we are interested chiefly in the heuristic methods which, in view of given conditions, thinking may use in finding the various solution-phases.

I experimented a great deal with the following problem: Why are all six-place numbers, of the form 267,276, 591,591, 112,112, divisible by 13? Here a single individual protocol may be given which, aside from a few typical as well as fruitless aberrations, contains the most practicable way of solution:

1. Are the triplets themselves perhaps divisible by 13?
2. Is there perhaps some sort of rule here about the sum of the digits, as there is with divisibility by 9?
3. The thing must follow from a hidden common principle of structure—the first triplet is 10 times the second, 591,591 is 591 multiplied by 11, no: by 101. (E: So?) No: by 1001. Is 1001 divisible by 13? (Total duration 14 minutes.)

For the present, let us consider only process 3. It begins with an *analysis of the goal.*[45] For the proposition that all numbers of the type abcabc are divisible by 13 means, on close examination, only that the divisibility by 13 is derivable from a property common to these numbers. The S now searches for such a common property, more precisely, for a "structural property", a common character relevant to divisibility. With this a process of *analysis of the situation,* more exactly of *analysis of premises,* is introduced.

Since what is required is characterized as "a common character relevant to divisibility", the search is restricted to a limited province. However, there are Ss who at this point seek only some common character other than that given. Consequently, they are often "stuck" in the purely visual. They observe, for instance, that the first and the last of any four consecutive digits are equal, which is of course of no further help.

Now the following is important: what is sought is not yet characterized as a "common divisor", but at most as "relevant to divisibility." The relation of ground and consequence which is decisive for the solution: "if a common divisor of numbers is divisible by q, then the numbers themselves are divisible by q," enters into the process only after a relatively vague analysis of the situation has already disclosed a part of the "if" premise, of the "ground." In the present case, this premise consists of two parts: (1) such numbers are divisible by 1001, and (2) 1001 is divisible by 13. And only by the discovery of part-premise 1 does the thinking process hit upon the decisive relation of ground and consequence. From this, the second part-premise is then organically derived, as its "completion" ("Is 1001 perhaps divisible by 13?"). Thus the discoveries of the

[45] "Goal" of course does not mean here "practical goal", as it did earlier, but that which is to be comprehended, which is to be proven. Strictly speaking, analysis of the goal is therefore here *"analysis of the proposition"*.

first and of the second part-premises occur under quite different conditions: the former precedes the decisive proposition, the latter is dictated by this proposition.

To 45 Ss (35 in three group experiments, 10 in individual, experiments), I gave this problem without any aid. The decisive ground-consequence relations was for none of them suggested directly by the problem's goal. Not one single time did the phrase: "I will just see whether the numbers have any divisor divisible by 13" arise directly from the original problem-setting. If it had, this demand would then have been imposed on a situation as yet unanalyzed.[46] On the other hand, every time the common divisor 1001 had been arrived at[47] by a process of analysis of premises—although not necessarily just that described above —the S immediately went on to investigate whether 1001 was divisible by 13. That is, the S experienced this inquiry as still to be made. Thus it becomes indirectly apparent that the decisive ground-consequence relation is now coming into play. Incidentally, all these Ss—at least at the question, "Why?"—explained their solutions by the existence of this "familiar" and "evident" relation.

This result, which will soon prove typical, may be formulated as follows: *The decisive ground-consequence relation which connects solution and goal is here "suggested from below."*[48] *It is suggested by a part of the premise, a part which is found through analysis of presuppositions and of goal.* The rest of the solution is found as *"completion"* of the ground-consequence relation so discovered.—In our example, the analysis of premises had the form of a "re-centering"[49] of the material originally given: abcabc = abc × 1001.

Let us here briefly discuss a few proposals which are fruitless, although often produced, and which stemmed chiefly from the fact that the Ss had not yet grasped the generality of the pattern abcabc, and consequently gave too much attention to the concrete examples. Part two of the individual protocol given above on page 31 raises the question of a rule about the sum of the digits. Behind the search for such a rule lies the general knowledge that a relevant relation sometimes exists between the sum of digits and divisibility.—The following wild explication of the premises was clearly inspired by the notion of the sum of the digits: "$2 + 7 + 6 = 5 + 9 + 1 = 15$ [cf. the numerical examples on page 31]. Is this significant?" Of 13 protocols more closely examined, 6 referred to this notion of summing digits. The idea in the single protocol mentioned above—that the pairs or triplets which constituted the examples might themselves be divisible by 13[50]—occurred with the same frequency (namely 6/13). —In 2/13 instances, the Ss investigated whether the examples might be powers of 13.—In 3/13 instances, the S specialized his model, e.g., he tried 100,100, because, as is quite well known, the general principle of a solution is often more evident in certain particular cases.

[46] By this I do not mean that some such thing could never occur, e.g., with practised mathematicians. We are interested only in the fact that the other way exists, and that, as is yet to be shown, it is a typical way of finding the solution.

[47] Among those 45 Ss, this happened with all who arrived at the solution, namely with nine. (The 35 Ss of the group experiments had only about five minutes for solving the problem.)

[48] Cf. Chapter I, 11.

[49] Max Wertheimer (45).

[50] From this it would of course follow that the whole six-place numbers as well are divisible by 13.

(Occasionally, such specializations appear also with other problems. In fact, they are often quite sensible and represent a rather general heuristic method.)

2. EXPERIMENTS WITH VARIOUS AIDS

In order to test the possible effect of certain hunches on the further course of the process, one can use the method of aids in the following way: the experimenter throws such ideas into the process from without, and observes what effects they typically have. With the "13" problem, I carried out group experiments in which six different groups of Ss received different aids during the process:

Aid a) "The numbers are divisible by 1001",
Aid b) "1001 is divisible by 13",
Aid c) "If a common divisor of numbers is divisible by 13, then they are all divisible by 13",
Aid d) "If a divisor of a number is divisible by *p*, then the number itself is divisible by *p*",
Aid e) "Different numbers can have in common a divisor which is in turn divisible",
Aid f) "Look for a more fundamental common character from which the divisibility by 13 becomes evident".

A seventh group (control group) received the problem without any aid (w.a.).

In other words, in both *a* and *b* a concrete premise of the decisive ground-consequence relation is given as aid; in *c* this latter itself is given (in a more general form, i.e., still lacking application); in both *d* and *e*, an abstract component of the decisive relation (in *d* by far the more important one) is given; and finally in *f* the allusion to explication of a more fundamental common character.—If one considers these six aids with reference to whether and in what genetic succession they appear spontaneously in a solution-process left to itself, one gets the following result. In 45 experiments set up *without aids,* the hunch *f* ("deeper common character") always occurred before all the others; *a* ("divisible by 1001") always before all except *f;* and *c* (the decisive ground-consequence relation) always before *b* ("1001 divisible by 13"). The variations *d* and *e* are probably not to be expected at all as spontaneous phases.

Table 1 contains all the collected data from three group experiments, namely the number of Ss (in %) who solved the problem in each of the seven groups.

TABLE 1

Groups	No. of Ss*	Percentage of Ss who solved the problem
a	22	59
b	10	50
c	13	15
d	22	14
e	10	0
f	13	15
w.a.	26	8

* In one of the three group experiments, the subjects were 63 students of the University of Berlin, who took part in a psychology seminar for beginners. In the two other group experiments, last year students of two secondary schools were concerned (53 Ss). In the data which interest us, all three experiments agree.

From Table 1 we may derive two important facts:

1. The premises containing the concrete allusion to 1001 (*a* and *b*) facilitate the solution far more than all the other aids. More particularly: aid *b,* which, as was mentioned, never appears as a spontaneous phase of a solution-process unless preceded by *a,* but which has in common with the latter the concrete allusion to 1001, helps about as much as *a*.

2. The other aids are of practically no help. (This applies also to *f,* perhaps because this phase hardly needs a special

aid.) For, the difference between 15% and 0% can hardly be assumed to exceed chance variation, as follows from the fact alone that the w.a. group betters the *e*-group by 8%. Aid *e* may not be an actual aid, but it is hardly to be described as a hindrance.

In these results, it seems to me worth noting that the pertinent general ground-consequence relation (cf. *c* or *d*), given as an aid, is of practically no help. From what was previously said, we knew only that, in the unaided process, it hardly ever arises unless there is some effort to analyse the premises. Now we see that, when artificially "grafted on," it hardly facilitates the solution, while its concrete premises, containing the reference to 1001, do so to a high degree.

3. THE REAL DIFFICULTY IN THE "13" PROBLEM

The report on the varying efficacy of various aids confirms an inference which, in fact, we might have drawn before: *the real difficulty of the 13 problem is overcome as soon as the common divisor 1001 emerges.* To average subjects, numbers of the type abcabc do not easily look like multiples of 1001, not even if the Ss are expressly urged to search for a common divisor (cf. aids *c, d, e*). Qualitative data confirm very impressively the difficulty of this "re-centering". For the most part, the Ss were much surprised when it was finally pointed out to them that the numbers were multiples of 1001 (had the form abc × 1001), which they then certainly had to admit. Before, they had merely applied the customary procedure of long division to the examples.

How firmly the banal structure deriving from mechanical division is impressed on the numbers, is shown as well by the following protocol. By analysis of the goal, a S had arrived at this idea, which is not at all bad: "Doesn't the proposition mean that the remainder from the division of abc by 13, set in the thousand-digit place before abc, must result in a four-place number divisible by 13?" But in spite of very great effort, the S got no further. And yet it would have been necessary only to apply this observation to the second triplet as well, and to relate the two remainders. The result would have been a number of the form *roor,* which would at once have revealed its relationship to 1001. But the penetrating realization was not reached that two or more remainders from the division of separate summands of a number together yield a summand of the same number. For the S concerned, the numbers were too "poor in relational aspects"[51] for this.

If the difficulty depends quite essentially on the fact that, in the structure abcabc as usually conceived, the divisor 1001 is well "hidden", it follows that the difficulty must decrease with every variation of the illustrative numbers which makes the divisor 1001 more evident and thereby more accessible to an explicatory analysis.—Such a variation was suggested to me by an individual experiment which took the following course: "If the proposition is valid, then any two consecutive numbers of the type abcabc must also be divisible by 13, e.g., 276,276 and 277,277[52]—and therefore also their difference, 1001. Now this will have to be proved. Is 1001 divisible by 13?—Yes."

In this process, we are interested in the decisive analytical step. The S had seen spontaneously that genuine "neighbors" exist among numbers of the type abcabc. This view of the situation shows some independence from the usual conception, and is already quite close to the required

[51] See page 39 and Chapter VIII.
[52] Analysis of goal.

structure: "276 thousands plus 276 units".—The "neighbor"-relation points to the constant difference and thus suggests the following ground-consequence relation, which is evident and also well known: "If a number a as well as the difference a-b is divisible by q, then b is also divisible by q". (It is clear that here again the decisive relation is reached only by way of an analysis of goal and of premises.)

Now to this solution-process I owe the following variation of the 13 problem, which was meant to facilitate the discovery of 1001. The text remained just the same except for the three examples. These were no longer "276,276; 591,591; 112,112," but, "276.276; 277,277; 278,278." The effect (see Table 2) was astounding. The "neighbor"-relation of the illustrative numbers set the situation-analysis on the trail of the difference 1001, and thus led to the solutions.

Although the numbers in Table 2 appear too small, they are not so, and for the following reasons: According to our former experiments with the old setting and on similar subject-material, 0.4 solutions are to be expected with five subjects, i.e., 0 corresponds entirely to the result to be expected. Furthermore, the fact that in the new setting 3 out of 4 Ss solved the problem corresponds to a percentage of solutions which I could otherwise approximate only by the most effective aids (for instance, "the numbers are divisible by 1001"). But above all, the three Ss actually reached the solution through 1001 as a *difference* between examples, a quite unusual procedure.

4. A MORE PRECISE DEMAND FACILITATES RESTRUCTURATION

We observed that a solution which involves a re-centering, a change of aspect, of the given material can be facilitated if the new aspect is somehow brought into prominence, and thereby made more accessible to an analysis of premises. This may be done by an explicit hint (cf. p. 33), or by suitable variation of the material (cf. p. 34). Now, it has already been pointed out in Chapter I, p. 11, that a solution always arises *"out of demands made by what is required on what is given"*. Consequently, a solution,

TABLE 2

Examples	Ss	Solutions
276,276 / 591,591	5	0
276,276 / 277,277	4	3

specifically a helpful restructuration of material, may be facilitated or hindered from two sides: not only by suitable variation of what is given ("from below"), but also by suitable variation of the *demand* ("from above"). In the 13 problem we helped from below, by bringing into prominence, by loosening up, the crucial aspect: "multiple of 1001". An attempt to help from above, in which the S was expressly urged to look for a common divisor divisible by 13 (cf. aids c and d in p. 33), proved unsuccessful —at least in our concrete instance. Now, it may not be inferred from this that a re-centering could never be facilitated from above, i.e., by "sharpening" the demand in a suitable fashion. Such a conclusion would be mistaken, as is shown by the following experiments on a "visualization" problem by G. Katona (18). A regular tetrahedron is to be put together out of its two congruent halves (see Fig. 11). Each half has as bounding surfaces two equal trapezoids, two equal triangles, and one square. The appearance of a tetrahedron is explained to the

S (sometimes, a whole tetrahedron is shown him as well). Thereupon the two halves, or only one, are given, i.e., first exhibited from all sides, to make apparent the surfaces of the parts and their congruence, and then set before the subject. The S may not attempt to put them

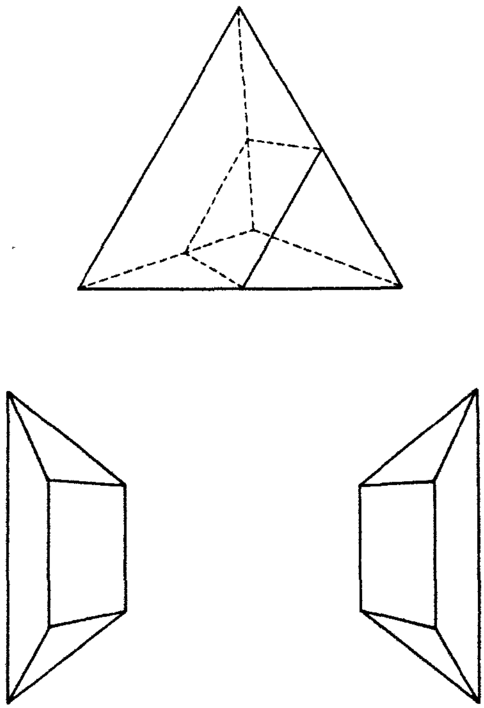


FIG. 11

together actually; he is to rely on thought.

The solution involves radical restructurings: surfaces and edges which were homologous in the two halves come into heterologous positions in the whole. For instance, one of the long edges becomes the edge of a side, the other of a base, etc.[53]

With seven Ss, these results were found: *all solutions finally arose by way of a problem-phase directed to forming triangles.* (Of these, four solutions required between two and five minutes, two between ten and twenty minutes, and one was called forth promptly by the aid, "set up a triangle".) Mostly, the attempt was made to supplement one of the trapezoids to form a triangle. Once the precise demand for triangular surfaces had arisen, the solution followed quickly, the restructuring occurred as if of its own accord, it "just had to happen".[54] The direction towards triangular surfaces had naturally arisen from an analysis of the goal. "Produce a tetrahedron" means, in any event, "produce triangular surfaces".

We can therefore say: *Generally speaking, a restructuring is the more easily carried out, the more sharply the new structure is "envisioned" in the demand.*

5. THE "ALTITUDES" PROBLEM

Many experiments were carried out on a second problem requiring a mathematical proof, this time in geometry. It runs as follows: "If the base-points of the three altitudes of a triangle are connected by straight lines the result is the triangle whose vertices lie on these base-points. Why do the altitudes bisect the angles of this triangle? Why, for exexample, $\alpha = \beta$?" (See Fig. 12.)

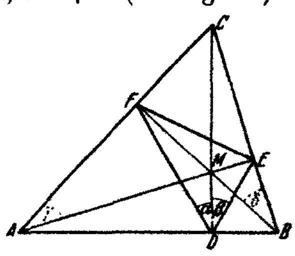


FIG. 12

[53] Katona explains the difficulty of the task principally by an optical illusion: the squares appear lengthened parallel to the longer edge. This illusion can be eliminated, and yet the task does not become much easier.

[54] In spite of this "theoretically" clear solution, two Ss could not visualize distinctly how the parts really combined to a whole tetrahedron. Somewhere there remained a gap.

A protocol: "Altitudes are given. The altitudes are perpendicular to the sides of ABC; thus $\angle\gamma$ and $\angle\delta$, with corresponding sides perpendicular to each other, are equal." (Analysis of premises.) "It must now be shown that $\alpha = \gamma$ and $\beta = \delta$."

The S got no further. After a few vain attempts to solve this last problem-phase, he gave up.—The process could perhaps proceed as follows: $\alpha = \gamma$, which means that α and γ are peripheral angles subtended by FM, that the points F, M, A, D therewith lie on a circle (analysis of goal).—Therefore it would have to be shown that F, M, A and D do in fact lie on a circle. Now by hypothesis, the angles at D and F are right angles subtended by the same chord AM[55]; therefore according to Thales' theorem, F, M, A and D do actually lie on a circle, ergo. . . .

Let us examine the process more closely.

1. First of all, an analysis of premises takes place. Such a process of making the given premises explicit[56] of course always occurs *sub specie of the goal.* As at an analogous point of the 13 problem a common character of the numbers was generally sought which was relevant to divisibility (not *any* common character), here an inference from the premises is sought which is relevant specifically to equality of angles. The Ss often expressly search for "angles related to α and β," or for "similar triangles in which α and β lie," and the like. It is clear that *the knowledge of very general relations*—here of the similarity theorems, of the congruence theorems, and the like—*gives direction to the analysis of premises.* From the most varied problems, I have protocols in which the Ss expressly ask themselves: *"What theorems are there about such a thing . . .?"* (An important heuristic method!) For example, with the altitudes problem, a S asked himself for "applicable theorems on equality of angles," whereupon only the similarity and congruence theorems occurred to him at first. When this would not do—the figure gave promise of no similar or congruent triangles—he asked himself further for "other applicable theorems . . ." and thus arrived at the peripheral angles theorem, which however he did not know how to apply. Another S actually did reach the solution by way of the theorem of peripheral angles.—But let us remember: What actually occurs to one, when he looks for theorems relevant to the goal, does not depend only on the store of knowledge *in abstracto.* The concrete way in which the model given (the "figure") is structured for the S plays an influential part.

2. After it had thus been discovered through analysis of premises that $\gamma = \delta$, a supplementary search was made to determine if $\alpha = \gamma$ and $\beta = \delta$. In the present problem, the decisive ground-consequece relation runs: "If two magnitudes are equal to two equal magnitudes, then they are equal to each other". Just as in the 13 problem, the ground-consequence relation here decisive is in itself extremely evident and familiar. But just as occurred there, the correspondence of the consequence (in this relation) with the general goal of the given problem is again for the most part not sufficient to suggest this ground-consequence relation. Rather, it enters the process typically only in the moment in which one concrete part-premise ($\gamma = \delta$) is revealed by explication. And just as happened there, the other part-premise ($\alpha = \gamma$ and

[55] The three altitudes intersect at one point.
[56] ["Making explicit" and "explication" are used in a parallel sense, explained on p. 43.]

$\beta = \delta$) is sought from this point as completion.[57]

6. GROUP EXPERIMENT WITH THE ALTITUDES PROBLEM

There follow the results of a group experiment in which almost all proposals consisted in deductions from the proposition, thus in *analyses of the goal.* (The final solution was not reached once by this means, although 26 Ss honestly strove it. To be sure, only about five minutes' time was granted them for a solution. Besides, the Ss, students of a *gymnasium,* were little trained in mathematics.)

The different proposals are given in the order of their frequency, which is in each case noted by a fraction whose denominator gives the number of Ss.

"If the proposition is valid, then the angles complementary to α and β must be equal. Perhaps this can be proven." 10/26

"If the proposition is valid, then the point of intersection of the altitudes, M, which is now the intersecting point of the angle-bisectors of DEF, must be at the same time the center of the inscribed circle of DEF. Now that would have to be proven." 5/26

"But then, the perpendiculars falling on the sides of DEF must (as radii) be of equal length. Are they?" 3/26

"If the proposition is valid, then the altitudes, as angle-bisectors, must intersect each other according to a definite relation." (False statement.) "What sort of relation is that? Isn't there a theorem of this sort?" 2/26

"If the proposition is valid, then the angle-bisectors must divide the opposite sides of the triangle in the relation of the adjacent sides. Can one prove that they do?" 1/26[58]

One recognizes that all these ways of solution represent explications of the goal. The second is not at all primitive. The Ss could hardly have suspected that it would finally come to nothing.

7. ON THE MANY-SIDED NATURE OF THOUGHT-MATERIAL AND ITS RESTRUCTURABILITY

A subject in an individual experiment (after trying many other ideas) hit upon an attempt similar to the second of the proposals just quoted. He even brought it—with a *tour de force*—to a solution. This is how it happened: "If the proposition is valid, then M, the point of intersection of the altitudes (in their character of angle bisectors), must at the same time be the center of the circumscribed circle of DEF.[59] For this to be true, M would also have to be the center of the inscribed circle of ABC, a circle which would have to pass through D,E and F.—Thus it would have to be shown, first, that M is the center of the inscribed circle of ABC.—This it is, as the point of intersection of the altitudes of ABC.[60] —Secondly, it would have to be shown that the inscribed circle of ABC passes through D, E and F.—Now, MD, ME and MF are perpendicular to the sides of ABC. Therefore the inscribed circle of

[57] This agreement between the two problems thus far given is not accidental. I was able to observe the same facts in still others of the twelve mathematical problems, deriving from quite different provinces, with which I have experimented. It seems that we have here hit upon *something* really typical.

[58] A few more explications of premises appeared as well. E.g., 2/26 Ss tried to explicate the figure with the help of the theorem of adjacent angles (applied to the triangle ADF). 2/26 would make use of the right-angled triangles.

[59] False proposition! The point of intersection of the angle-bisectors is the center of the *inscribed* circle.

[60] Here the difficulty is again met by a false proposition. The point of intersection of the altitudes is by no means the center of the inscribed circle.

ABC passes through D, E, and F, since the radii are perpendicular to the tangents."

In this process it is interesting, among other things, how knowledge which is no longer sufficiently definite is "bent into suitable form" for a solution. The mention of "angle-bisectors" made the S think of "inscribed or circumscribed circles". He saw at once that the circumscribed circle of DEF could function as the inscribed circle of ABC; then from this attribute, conclusions could perhaps be drawn referring back to the decisive attribute of being a circumscribed circle. In other words, he hoped that the angle-bisectors might lead to the inscribed circle of ABC by way of the circumscribed circle of DEF. It would be gratifying if the altitudes would sanction this attribute of being an inscribed circle, i.e., if their point of intersection were the center of the inscribed circle. In short, the two mistaken postulates did not simply "occur," but they functioned in a larger plan. This plan was aroused by the fact that angle-bisectors suggest inscribed or circumscribed circles, and by the further obvious fact that DEF and ABC were respectively inscribed and circumscribed.

The whole process, which was just "unfolded" in such detail, actually lasted only a few seconds. Although his geometrical knowledge was only rudimentary, the subject was extraordinarily gifted in this field. He had an excellent "perspective," as I should like to call it; he saw not only one definite aspect of the thing at a given time, but—through it, as it were—still further aspects. *Now since each aspect of the premises, in and for itself, could be the origin of some solution—and several such aspects dealt with successively that much more—the subject must be just so much better off, the more and the more varied aspects he is able to command at one glance,* i.e., *without tedious "work of explication."* Just so, the man whose vision is not limited to the few feet just ahead, but who directly takes in the more distant possibilities as well, will most surely and quickly find a practicable path through difficult terrain. The circumstances involved are similar to those we discussed a little earlier, from the negative side. There we saw that, for the solution of the 13 problem, the common (banal) aspect of a six-place number is useless, that another, more fundamental aspect is demanded. *The onesidedness, the poverty of aspect of thought-material is* —as we shall in every case distinctly observe—*the chief characteristic of poor thinking.*[61] A theory of individual differences, e.g., a theory of what mathematical talent actually is, must start from this. It must explain the origin of this poverty of aspects. For, to mention this in advance, insufficient acquaintance and lack of familiarity with pertinent general aspects does not explain everything.

In this connection, I should like to describe briefly yet another process—the only one, by the way, in which any of my subjects really succeeded in solving the altitudes problem.[62] After the S had first attempted a solution with similar triangles, then had searched for an indirect proof, he considered "peripheral angles in circles" for a moment and finally explicated—with no very definite direction—the right angles at D and F. With this, Thales' theorem occurred to him. $\angle ADM$ and $\angle AFM$ are right angles over the chord AM; furthermore, A, D, M and F lie on a circle; therefore the

[61] See Chapter VIII.

[62] Besides the group experiment mentioned above, I set the problem to five subjects individually.

theorem of peripheral angles is applicable.

That the theorem of peripheral angles occurred so early to this S, and then Thales' theorem, both of which are about circles, *that* is what is unusual. It does not require much to arrive at the inscribed circle from the angle-bisectors —with many Ss, "angle-bisectors" suggested the pertinent theorem—but to arrive at *the* circle by *another* way is really an accomplishment with this *figure so little suggestive of a circle.* One S said, when I had told him the solution, "I should never have come upon circles in connection with this figure." Clearly, the theorems are suggested[63] by certain pertinent aspects of the model which are reminiscent of them—of course, in conjunction with the goal.

8. WORK OF EXPLICATION UNTIL THE DECISIVE GROUND-CONSEQUENCE RELATION "SNAPS INTO PLACE"

We now see under what circumstances a solution will be especially difficult. *A solution will be the more difficult to find, the more work of explication it presupposes.* The whole solution-process may be divided into two stages:

1. The stage of explications (of the goal and especially of the premises) which should lead to the decisive ground-consequence relation, and must therefore themselves occur relatively at random.

2. The stage in which the decisive relation "snaps into place", with all the "completions" it makes necessary.[64]

[63] Compare the suggestion of the "divisor theorem" by the aspect: "the numbers are divisible by 1001" (p. 31), or the suggestion of the theorem: "If two magnitudes are equal to a third, then . . . ," by the aspect $\alpha = \delta$ (p. 37).

[64] These completions in their turn may of course be processes of the same type as the whole process, and may therefore also occur in these two stages.

Now, the longer is the stage of explication, the more difficult the task. As indicated, this "length" is a function not only of the objective number of the steps involved, but also a function of the subjective "transparency" and wealth of aspect of the material.

An example of a process which demands very much explication. The problem is to prove that "if $a.b$ is divisible by a prime number p, then either a or b is divisible by p." Presupposed are the concept of a common multiple of several numbers, i.e., of a number which is divisible by the numbers concerned, and further the concept of the smallest common multiple. Among many other propositions, the following proposition P is also supposed to be known: Every common multiple of numbers is a multiple of the smallest common multiple.

The process runs something like this: "Either a or b is divisible by p" means the same as "If p does not divide a, it divides b, i.e. $b = qp$". This is therefore to be demonstrated.—Now, if p does not divide a, $p.a$ is the smallest common multiple of p and a. Further, $a.b$ is a common multiple of a and p, therefore $a.b = q(p.a)$—according to proposition P. Therefore $b = q.p$, i.e., if p does not divide a, it does divide b. Q.E.D.

This process is interesting in several respects.

1. It begins by an explication of the goal which is completely evident, so evident and general that *once* it is achieved, one is inclined to consider it banal. In reality, this explication of the goal contains the principle of the whole process: to conclude from the assumed invalidity of the "either" clause the validity of the "or" clause—and this is anything but banal, for it is almost never carried out.[65]

[65] I set the above problem to three Ss, two of

2. That "p divides b" has the aspect "$b = q.p$" is familiar to any mathematician, and understandable by any layman. But to the layman it is at first a discovery.

3. *It appears thoroughly plausible that* if p does not divide a, $p.a$ is the smallest common multiple of p and a. And yet this might be one of the explications which already lie beyond the spontaneous perspective of one who is not acquainted with the theory of numbers, and which therefore requires of him a detailed deliberation *ad hoc.* Now consider: so early an explication as this can arise only more or less by chance, i.e., without knowledge of the If and Why of its later functioning. *But the more fortuitous an explication, the less difficulty can be put in its way if it is to occur at all; the more explicit, the "looser" must therefore be the aspects per se which are to be explicated. In order to divine ways and to test their applicability, one should not be forced to grope along each one clumsily as though one were shortsighted.*

4. That $p.a$ is the smallest common multiple of p and a actually does not yet lead to proper connection with the goal. Therefore further explications are needed. The next explication ($a.b$ as the common multiple of a and p) is already less blind or fortuitous; for at least it is subject to the more precise demand for an explication bearing on p and a. It may be that even the proposition P (p. 40) was already suggested, and that therefore a common multiple of p and a was actually being *sought for* (as "completion" of the P-relation). But even upon so precise and limited a demand, the discovery that $a.b$ is a common multiple of a and p is no small matter. The non-mathematician is disinclined to conceive $a.b$ as the common multiple of two numbers which had so far functioned in such different rôles.

5. Even with this last explication, the S still could not easily divine that the statement would follow from it directly (by cancellation). In other words, the explication stage of the present process extends to the last step but one. Only here does the "snapping into place" occur.

whom were mathematicians. None arrived at the aforementioned goal-explication during the experiment; all sorts of fruitless explications were undertaken instead. One S took the problem home with him, and after a few days brought me the proof above.

9. "THE FIGURE REMINDS OF . . ."

Those solutions are especially difficult which can be reached only by a "dodge," i.e., those in which the necessary explications lie in a direction attainable only by an extremely rare re-centering of the given facts. To these belongs the customary proof that the three altitudes of a triangle intersect at one point. Another example: It is to be shown that ACB ($\angle$gCA = $\angle$gCB) is shorter than any other connection between A and B which touches on the straight line g, e.g., than ADB (see Fig. 13). "The proof is based on a method which appears, from the purely mathematical standpoint, as a

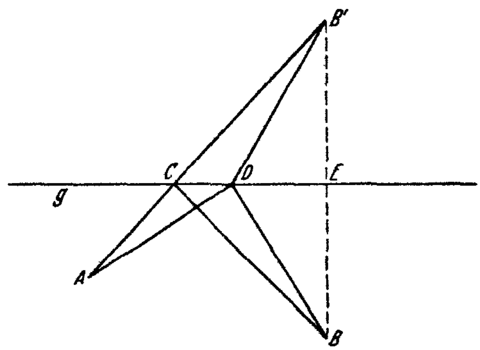


FIG. 13

mere trick, but which is suggested by the optical interpretation."[66]

This proof cannot be found unless the figure spontaneously reminds the subject of reflection in a mirror (which, however, did not happen a single time in my experiments with seven subjects talented in geometry). If it does, then a definite direction, otherwise very remote, is prescribed for the explication of the premises, and the solution can be found. If it does not, there is little likelihood that any lines will be extended to the other side of the straight line *g*.

Actually, Ss who were neither by the figure nor by the experimenter reminded of mirroring, explicated all sorts of things, but throughout "on this side" of the straight line. For instance, the line AB was drawn, so that triangles resulted, and perpendicular lines were dropped from C and D to AB; the surface-areas of the triangles were considered, etc.—The following enterprise (see Fig. 14) is

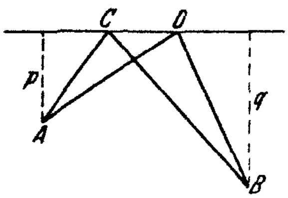


FIG. 14

quite clear-sighted: AC, CB, AD and DB were each expressed trigonometrically, i.e., as functions of the angles at C and D and of the perpendiculars p and q. The sums AC + CB and AD + DB were then compared, which, however, was finally of no further help.—All these explications take place on this side of *g*.

[66] Rademacher and Töplitz, (31), p. 20. One "reflects" the point B and the lines BC and BD in *g*. Because of the reflection, the triangles CBE and CB'E are congruent, therefore ∠BCE = ∠B'CE, therefore also ∠B'CE = ∠gCA, i.e., they are opposite angles. Therefore ACB' is in a straight line connecting A and B', and is consequently shorter than ADB'.

One solution is very interesting which actually arose, in principle, as the solution outlined above ought to have occurred. The figure reminded one of my Ss of the principle of the ellipse: an ellipse defined as the geometrical locus of all points the sum of whose distances from two fixed points (A and B) is constant and smaller than the sum of their distances from points outside. The proposition to be proven would mean, therefore—in terms of an ellipse—that the ellipse constructed with the sum AC + BC must touch the straight line *g* at C (explication of the goal); or—otherwise expressed—that *g* is the tangent to the ellipse at point C. When the S had come this far, he said, "Now I should need only the statement: the tangent to the ellipse forms two equal angles with the corresponding radii vectores (see AC and CB)." With this statement, the proof would indeed be achieved, since it had been assumed that *g* formed two equal angles with Ac and CB, and since there exists only *one* straight line which forms equal angles with AC and CB.

In this solution-process we have a nice illustration of what is meant by the phrase, "the figure reminds of . . ." *The figure before the subject suggests the "sphere" of certain relevant theorems,* and thereby stimulates these theorems themselves—an excellent example for the "heuristic function of inspection"[67] (see Chap. VIII, p. 108f.)

10. ANALYSIS OF GOAL AND SITUATION AS EXPLICATION OF PROPOSITION OR PREMISE

The experimental results of the present chapter have familiarized us with important heuristic methods of thinking. What we have already met in the pre-

[67] [*"Anschauung"*.]

ceding chapter under the name of analysis of situation and of goal, plays—in related forms—a rôle hardly to be overestimated in the solution of mathematical problems.

There is, in the first place, the analysis of goal in its form of explication of the proposition. Decisive solution-phases often arise through one's asking oneself, "What is really meant by what is to be shown?—How else could it be formulated?—What follows from the proposition, from which again this proposition itself could be proved?" *In other words, decisive solution-phases often arise as inferences from the proposition.* They are "re-centerings" of the goal, of the proposition—therefore not to be confused with re-centering of the problem-object, as the term occurs in Wertheimer's examples (45). They do not answer the question: "How must the object be conceived, so that the proposition can be proven from it?" Rather they answer the question: "How must the proposition be understood, so that it can be proved to hold for the object?"

This formulation of the question makes it clear that not *just any* equivalent inference from the proposition can be regarded as an explication of the goal. For, an equivalent inference from the asserted infinity of the series of prime numbers would be not only the productive assertion that for every prime number a greater one exists, but also something like this: "Mr. Smith is therefore mistaken if he believes that only a finite series of prime numbers exists". Or, to the proposition $\alpha = \beta$ (see p. 36), $\frac{\alpha}{2} = \frac{\beta}{2}$ would also be equivalent. From such irrelevant explications of the goal, no intelligent human being would expect to obtain the proof of the corresponding propositions. In a solution-process, an explication of the goal has the function of transforming the original proposition, which is as such "averse to proof", into one as nearly as possible demonstrable by the given means or premises. *The analyses of the goal, therefore, take place sub specie of the actual conditions of the problem*—exactly as Wertheimer's re-centerings take place *sub specie* of the proposition, of the goal.

By the way, explication of the goal is a procedure quite analogous to *analysis situs* as it is used in schools in connection with geometrical problems. Just as there one regards the construction as already complete, in order to read off[68] relevant relationships from it and to investigate their constructibility, just so, in explications of goals, the proposition is considered as demonstrated, and certain conclusions are derived from it whose demonstrability is then investigated.

The second method, which everywhere supplements the first, is the *analysis of the situation*—in the form of explication of premises.[69] With mathematical solution-processes, several different types of analyses of the situation occur:

1. Inferences from the premises (compare the mathematics instructor's admonition, "Make use of the premises").
2. Re-centering of the problem-object. This is Wertheimer's type of re-centering. It may be regarded as a reversible inference from the given conditions, or better, as leading to a problem-object equivalent to the original one. Therefore it is an exact counterpart of the explication of the goal, which produces an equivalent goal.
3. Simple observation of the figure, which visibly embodies the premises.
4. Suggestion of pertinent theorems by the figure.
5. Utilization of the available possibilities in the sense of possible supplementary lines.

[68] Cf. page 52, footnote 84.
[69] In the form of "analysis of conflict" it is somewhat less prominent in mathematical problems.

Explications of premises do not occur at random; there is no attempt to seize upon just any inference from the premises. Just as in explications of the goal, this is evident from the fact that countless explications of the premises are conceivable which do not appear and are surely not taken seriously in any solution-process unless it be a case of "flight of ideas." For example, in the 13 problem it is not explicated that the sum of the digits of a six-place number cannot exceed 54, and the like.—In short, *analyses of the situation also take place sub specie of the problem as a whole, specifically of the proposition, of the goal.* Their range of possibilities is limited by the whole.

Incidentally, it is probably quite clear that each concrete explication of the goal already involves analysis of premises—for the simple reason that every concrete goal already involves concrete premises. Naturally, it is otherwise with such direct explications of goals as: "either *a* or *b*" is tantamount to "if not *a*, then *b*."

Further, but much more specific, heuristic methods in solving mathematical problems would be, for example, starting from some suitable special case, making an illustrative sketch, searching for an indirect proof, and so forth.

11. DIDACTIC CONSEQUENCES

If I wish to demonstrate a proof to someone, I can proceed in quite different ways, according to whether I start more from what is given or more from what is demanded. The extremes are:

A. I try—so far as it is possible—to derive the proof from above, from the goal, from the proposition, by asking, "From what would the proposition follow? What is necessary to it?"

B. Or I begin from below and ask, "What is given?", i.e., I derive from the given premises various inferences, which later 'surprisingly' crystallize into the proof of the proposition.

I should like to name the first way the "organic" one. Here, from the "function" (functional value) arises the embodying material, the instrumental means.—The second and opposite way is the "mechanical" one. This is what Schopenhauer meant when he complained of trick proofs ("Mausefallenbeweise").[70]

Why "organic?" Why "mechanical?" The "organic" origin of the eye is that part of the skin which is sensitive to light; that of the steam engine is the clattering lid of the teakettle, the displacement of solid bodies by steam. To this idea, everything further is hierarchically subordinated as elaboration and differentiation.—It would have been a "mechanical" procedure if Nature had put the eye together out of its many highly-developed parts and auxiliary contrivances, or if man had "discovered" the steam engine by appropriate combination of pistons, cylinders, valves, etc.

So far as it is at all possible, mathematical instruction should proceed organically. To be sure, it may be thrilling when out of the darkness of long preparations the spark of the proof flashes forth—but it is unnatural, i.e., alien to the natural genesis of the new. Certainly, as we saw, one frequently has to start "from below" in a proof, but—as we saw at the same time—even such explications of premises need not happen without regard to what is demanded. In every productive explication of premises, there are certain "organic" phases, which at least limit the range.

The difference between a procedure as organic as possible and an unnecessarily

[70] Cf. "Ueber die vierfache Wurzel des Satzes vom zureichenden Grunde", #39.

mechanical procedure is now to be demonstrated *ad oculos* in an especially simple mathematical example.

Proposition: The point of intersection of the three perpendicular bisectors of a triangle is the center of the circumscribed circle. (See Fig. 15)

Proof:

Organic:

What does "center of the circumscribed circle" really mean? Clearly, the point equidistant from the three corners, therefore MA = MB = MC. That has to be proven.—MA and MB are sides of the triangles MAD and MBD. Thus, if possible, it is to be shown that these triangles are congruent.

In point of fact:
AD = BD
Further: ∠MDA = ∠MDB = 90° (by hypothesis)
Further: MD = MD, therefore the triangles are congruent (Congruence Theorem I), therefore MA = MB, etc.

Mechanical:

I connect M, the point of intersection of the three perpendicular bisectors, with A, B and C and consider, at first, the triangles MAD and MBD.

Here AD = BD
Further: ∠MDA = ∠MDB = 90° (by hypothesis)
Further: MD = MD, therefore the triangles are congruent (Congruence Theorem I), therefore MA = MB, etc.

But this means that M is the center of the circumscribed circle. Q.E.D.

We can see that the way of proof on the right is simply the reverse of that on the left. The mechanical way of proof is certainly shorter, more economical to set down. But the organic is more closely related to the way in which a new productive performance naturally arises—at least in principle, for surely, with so simple a proof, both ways are about equally "natural."

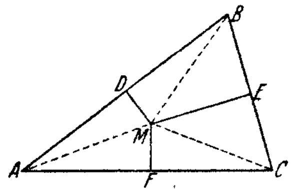


FIG. 15

To summarize: In teaching, it is advisable—be it even at the cost of brevity and "elegance"—to proceed as organically as is feasible, to undertake as few fortuitous explications of premises as possible.

PART TWO

INSIGHT, LEARNING AND SIMPLE FINDING

CHAPTER IV

ON TOTAL INSIGHT OR EVIDENCE

1. THE GRASPING OF GROUND-CONSEQUENCE RELATIONS AS NECESSARY TO PROBLEM-SOLVING

THE PAST three chapters have already made sundry contributions towards answering the main question we are investigating: How does the solution arise out of a *Problemstellung?* We found, firstly, that the final solution is mediated by successive reformulations of the problem, and secondly, that these reformulations or solution-phases are in their turn mediated by general heuristic methods.

But this is still not a complete answer. As happens particularly with mathematical problems, the method of solution may consist, for example, in making appropriate deductions from the proposition or from the premises. But then the psychologist must still discover the psychological meaning of "making deductions from", must discover how, from one fact, thinking actually brings about the intelligent transition to another, new fact. Or the method of solution may consist in examining the situation for flexible conflict-elements—as happens especially with practical problems. Yet this observation in no way reveals how such a conflict-element makes itself known to thinking *as ground or cause.*[71] To be a conflict-element means of course to be ground or cause of the conflict. And analogously, to be a solution means to be ground or cause of the realized goal. To understand a solution as solution accordingly means to understand the solution as ground of the goal.—But with this, analysis of conflict and grasping of the solution are involved in the whole set of problems concerned with the apprehension of causation. How does thinking succeed in reading off the cause from an effect, or the effect from a cause?

These two questions, namely, how thinking is able from the logical ground to apprehend the logical consequence and from the cause to apprehend the effect, we shall combine in one question and give the following generalized formulation: *Of what nature are the ground-consequence relations which are important for thinking in problem-solving?* Or, otherwise expressed: What possibilities has thinking of obtaining, from the nature of a ground, information on the nature of the consequence? The next two chapters will be dedicated to this investigation, which is fundamental no less to the psychology of thinking than to philosophy. (In this connection, something will also have to be said on the rational nature and the genesis of "heuristic methods.")

2. DEFINITIONS

A connection of two data a *and* b *may be called "totally intelligible", if it can be directly understood from* a *that, if* a *is valid, then* b *and precisely* b *is valid. (A connection is therefore "unintelligible" to the extent to which it is to be "accepted as mere fact".) A connection may be called "partially intelligible" if*

[71] In Chapter II, 6 and 7, we investigated only the disposability, the "looseness" of a conflict-element, the condition of its being "knowable," not of its "being known".

at least certain features of b *can be understood from* a*—or are at least singled out by* a *in contradistinction to other possibilities.* In other words, a *b* may be called intelligible in respect to an *a,* if and in the degree to which *b*'s real phenomenal characteristics, in contrast to other possible characteristics, appear *directly favored* by the phenomenal nature of *a.* By 'directly' I mean: without intervention of further factors.—In the present chapter, only *total* intelligibility is to be dealt with.

The definition given above of intelligibility has reference to a fundamental statement from Hume's "Inquiry Concerning Human Understanding" (Sec. 7, Part 2): "When any natural object or event is presented, it is impossible for us, by any sagacity or penetration, to discover, or even to conjecture, without experience, what event will result from it . . ."—But suppose that, in regard to its characteristics, *b* is singled out by *a* in contrast to other possibilities. To the degree to which this occurs, one must also be able "to discover or even to conjecture" the *b* directly from the *a,* be it wholly or in part.

As prototype of entirely unintelligible connections we may cite one of those if-then relations which Thorndike imposed on his cats (39). If the cat licks itself, then the cage door opens, or if it presses on a certain knob, then the same happens. From the circumstance—to be understood quite literally—that a cat licks itself or presses on a certain knob, it appears in no way intelligible that the cage door should open.—Just as in the field of causation, we can find any number of cases of such unintelligibility in the province of coexistence. From the fact that a house belongs to a Mr. N, or that it has so and so many windows, it is in no way evident that it has a gray roof. (Be it noted that, in both examples, the premises should mean only what is expressly stated in them. Thus, for example, "Mr. N" is not an abbreviation for a definite concrete man with such and such characteristics, but means only: some man, purely and simply.)

In contrast to this it undoubtedly does happen that *b "follows"* from *a,* i.e., that *b* is completely intelligible from the content of *a.* For example, it follows from the circumstance: "the house is higher than the tree and the tree higher than the bush," that "the house is higher than the bush."—How is such intelligible connection of ground and consequence, such evidence, possible?

3. ON ANALYTIC AND SYNTHETIC EXPLICABILITY[72]

Two classic answers to this question are: Either *b* is already co-contained in *a* and may therefore be "analytically" explicated from *a,* or the faculty of reasoning is so constituted that it must always and everywhere connect *a* with *b.* The second formula, with which Kant believed he must and could insure the possibility of empirical knowledge, need not occupy us further here. It does not fully correspond to our definition of the intelligible, which required that it should be evident from *a* that, if *a,* then *b.* Simple necessity or generality of the connection is not sufficient; *b* must be evident "from the essence of *a*"—as Husserl would have put it.

The first answer, on the other hand, is to be inspected more closely. What does it mean to be "co-contained?" Something can be co-contained: 1) *as a constituent of the whole,* i.e., *so that it follows from the whole, but not from the other parts alone (constitutive co-con-*

[72] [*Ablesbarkeit.*]

tainedness).[73] Examples for this would be: a) (constitutive co-implication)[74] "The roan is reddish," or, "A straight line is determined by two points" (this statement understood as axiomatic *definition* of the straight line). Here the predicate, the reddish coloration or the being determined by two points, is constitutively co-contained in the subject-concept defined by the statement, and indeed "co-implied."—b) (Constitutive co-existence)[75] "That house yonder has a gray roof," or, "The knowledge of the purpose and of the means belongs to the essence of action." Here the predicate, the gray color of the roof or the knowledge of . . . , is constitutively co-contained, "co-existent" in the full concreteness of the given.[76] Upon the fact that the parts and elements are given as such parts within a concrete experienced whole depends the usual evidence of inspection,[77] the "explicability" of the parts from the completely given whole.

One sees that constitutive co-containedness undoubtedly makes possible an evident "following" of the parts from the whole, that without doubt the conclusion here arises wholly intelligibly from the premises. But if no other connection between ground and consequence were intelligible to thinking than this analysis on the basis of a *constitutive co-containedness,* such thinking would be of little use in the world. Our question: "How is intelligible connection of ground and consequence possible?" calls for another answer.

And in fact, something can be co-contained (if one still wishes so to call it): 2) *as a consequence of the other elements of the whole (non-constitutive co-containedness)*.[78] How about the example: "From $a > b > c$ it follows that $a > c$?" I can demonstrate the meaning of 'greater' on all kinds of pairs of objects. With the help of the concept so obtained, a clear "paradigmatic"[79] situation: "*a* greater than *b* and *b* greater than *c*," can be constructed. From the situation so constructed, the fact: "*a* greater than *c*" may now be *read off. Yet —and this is the important point—this fact was not needed in the construction of the situation, the "foundation" from which it is now being read off. Therefore, in this case the conclusion is not constitutively co-contained in the premises.* (On the other hand, it would be reading-off on the basis of a constitutive co-containedness if, for example, "$b > c$" were explicated from the statement: "$a > b$ and $b > c$," or if from the concept "greater than," already defined as transitive, the transitiveness were explicated.) The reading-off of a fact which is *non-constitutively* co-contained we shall call *"synthetic explication."* In this, a new aspect is "affixed" to the situation constituted in the premises (compare Kant's expression, "synthetic judgment"). The "extraction" of a *constitutively* co-contained element, on the other hand, may be called "analytic."

Another example, more or less of the same type: Without the use of arithmetic

[73] [*Konstitutives Mitenthaltensein.*]

[74] [*Konstitutive Mitgesetztheit.*]

[75] [*Konstitutive Mitgegebenheit.*]

[76] In the last example it appears as if the predicated element already followed from the other elements (namely of action). This apparent following from the other elements always exists if a structure, a "strong figure" (Köhler), is present, in which all elements influence each other mutually. It is clear that then every element "follows" from these its "imprints" in the other elements. Nevertheless, each element can be entirely constitutive or indispensable for the whole.

[77] [*Anschauung.*]

[78] [*Nicht-konstitutives Mitenthaltensein.*]

[79] We shall call a situation "paradigmatic" if it is constructed for the mind's eye by the exclusive use of the concepts expressly contained in the premises.

axioms, it can be demonstrated that a divisor of a divisor of the number a is a divisor of a. This may be done in the following way: a "divisor" of a whole G may be geometrically conceived (defined) as a magnitude T whose finite repetition exactly fills out the whole, as it is paradigmatically represented in Fig. 16a. A divisor T′ of T is related to T exactly as is the latter to G (Fig. 16b). Then I see,

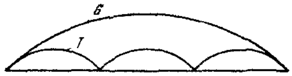


FIG. 16a

FIG. 16b

I read off, that the finite repetition of T′ produces G, i.e., that T′ is a divisor of G., Q.E.D. Here also, the conclusion has not been utilized in constructing the original situation (or—what is equivalent to this—for the definition of the premise-concepts). Thus there is no constitutive co-containedness here.

A third example: I try to have two paradigmatic "straight lines" intersect twice. Then I see, I read off, that in this effort they have become bent under my hands; more precisely: the being bent proves to be a new aspect, i.e., not utilized for construction of the situation: two lines intersect twice. With this, in conjunction with the principle of contradiction, the axiom: "Two straight lines intersect no more than once" is indirectly made evident.[80]

From these few examples it is already clear that, according to their direct meaning, the concepts "analytic" and "synthetic" are not unconditionally opposite, but are two sides of a definite relation. For, even "synthetic" reading-off is based on a definite form of co-containedness, and is therefore—in the widest sense—"analysis." The term "analytic" has, however, a special affinity to constitutive co-containedness, and the term "synthetic" to non-constitutive co-containedness.

A comment on the concept of "inspection" used here: It makes no difference whether one defines what is here named "inspection" as inspection in a narrower sense, or in part as "representation without inspection of contents" (compare the so-called "imageless thought"). "Inspection" will be to us that sort of event in which something like "$a > b$, $b > c$, $a < c$" can*not* be accomplished. It might coincide with Husserl's "intuition," in contrast to pure "signification," mere "meaning."[81] Incidentally, Husserl calls such perception as that of $a > b$ "categorical inspection" [kategoriale (fundierte) Anschauung].[82]

4. DISCUSSION OF MODERN POSTULATIONAL THEORY

A concise discussion of modern mathematics should not longer be postponed. What attitude do the proponents of modern postulational theory take to the question of insight or evidence? For the modern mathematician (12), the straight line is that entity—originally undefined—which receives definition from the axioms about it, nothing more. And the relation: "greater than" acquires its transitiveness in mathematics merely by logical deduction from certain postulates, from the axioms about the relations: "between" and "congruent." In other words, in modern mathematics, propositions—re-

[80] Of course, this could also have been done in the form of a direct proof: from the construction of two straight lines which intersect once, it may be read off that they do not meet again.

[81] See especially *15*, II, p. 192.
[82] *15*, II, p. 128 ff.

lations of the form: "if a, then b"—which are not logically deduced from others are postulates, conventions (Poincaré), implicit definitions (Schlick). The evidence of these axioms is thus, intentionally, reduced to the form of a constitutive co-containedness (more specifically, co-implication) of b in the defined a, in short, to "analytical" explicability (tautology). Metaphorically speaking, the mathematician prescribes their "evidence" to the mathematical objects, thus sacrificing synthetic evidence. (By the way, nowadays the *logical* axioms, the principles of deduction, are also preferably regarded as mere postulates.)

There is no objection to this procedure. The mathematician and logician may proceed thus. As a matter of fact, mathematics and logic would "function" even if all principles were nothing but postulates. In short, "postulation" and "evident following from" are equivalent in respect to their logical achievements.

However, even though certain if-then relations *may* be treated as mere postulates, they can also be seen in a different light. Perhaps this other conception is irrelevant to the *purely logical* structure of mathematics. It nevertheless corresponds to actual facts in thinking. For, this other aspect of such if-then relations consists in precisely that synthetic evidence which we tried to demonstrate above with a few examples (p. 49 f.). Wherever mere postulates rule today, older forms of mathematical procedure demanded that such evidence be given. The demand has been relinquished by modern mathematics. Thus real evidence has been banished from the province of "pure" mathematics. To be sure, meanwhile evidence has acquired great honor in philosophy as the principle of so-called phenomenological knowledge, and has in this rôle been sharply contrasted with all merely inductive certainty. But if with reference to such knowledge the phenomenologists like to use the expression "im Wesen gründen," they only hide with this expression the following fundamental problem: how can evidence arise from mere inspection of a subject matter?

Incidentally, it makes no difference in principle to the *applicability* of mathematical axioms to *empirical reality*, whether the axioms are practical postulates or instances of evidence. For, if I abstract from what is empirically given certain concepts such as "greater than" or "straight line," if I combine them into more complex structures, and then make all kinds of new evident explications from them, I have no better guarantee that these explications will coincide with reality than if I dealt only with practical postulates. I might after all have erred in the abstraction. For example, short sections of slightly curved lines seem straight to the eye. Consequently, I can abstract from reality the concept of the infinitely extended straight line, and believe it to be realized;[83] and yet there need not be anything of the sort in this same reality. But from this it follows that, even for concepts which are abstracted from reality, applicability to this reality is in no way obligatory. Hence, no so-called conclusions from essence (*Wesensschlüsse*) apply directly and necessarily to reality. *For whether the reality concerned actually embodies the "essence" expressed in the premises or not, can be proven in no way but empirically.* Einstein's statement is valid for synthetic just as for analytic intelligibility: "Insofar as the propositions of mathematics refer to reality, they are uncer-

[83] By the way, it could be abstracted even from clearly curved lines.

tain, and insofar as they are certain, they do not refer to reality." (5, p. 3 f.)

If, as we saw, neither pure mathematics nor the investigation of reality relies upon synthetic evidence, one could ask why synthetic evidence is necessary at all. The answer is threefold: 1) Whether necessary or not, it exists. 2) Even a maximally "formalized" mathematics still contains certain basic intuitions which cannot be eliminated, e.g., the concepts "element," "relation," "following from," as well as the principles of deduction. Should not what is right in the case of *logical* inspection or evidence be just as right in the case of *spatial* inspection and evidence? Further, if one believes that spatial inspection is unable to assure itself once for all and *per evidentiam* of its correspondence with the axioms of Euclidean geometry, then "pure spatial inspection" would be placed in the same epistemological position as "empirical reality." This appears to me extraordinarily erroneous, despite Reichenbach's attempts to this end (32). 3) Most important: As we shall see, without synthetic insight and evidence, productive thinking is nowhere *psychologically possible,* either in mathematics and logic or in the investigation of reality. More generally formulated: *The kind of experience in which synthetic evidence occurs represents the psychological medium of all productive thinking, postulational thinking included.* It is with this relevance of evidence for actual thought that we are here concerned.

5. HOW IS SYNTHETIC EVIDENCE POSSIBLE?

Yet before we approach the proof for this more general proposition, something still remains to be done. Our question: "How is synthetic evidence possible?" has not yet been fully answered. Now, obviously—if we compare the examples on p. 49 f.—*synthetic evidence is possible for this reason. As a rule a situation may be constructed, or defined, by means of fewer facts than can afterwards be read off from it if new points of view are applied.*[84] *Not all possible aspects of a thought-object are necessary to its construction, just as little as all possible aspects of a visual object are necessary to the unambiguous comprehension of its structure. One and the same situation, paradigmatically constructed for inspection, may be considered a) from new sides, b) in new directions, c) in new structurings, d) for the first time as a whole, etc.* It is this "aspect structure" of the thought-objects which makes synthetic evidence possible.

Examples for

a. the necessary curving of two "straight lines" which intersect more than once,

b. from "*a* greater than *b*" follows "*b* less than *a*,"

c. the transitiveness of the relations "greater than" and "divisor of,"

d. if there are in a room a person A and a person B and a person C . . . then there are *three* persons in the room.

Such aspects as are in this way explicable from a situation constructed heterogeneously, i.e., by means of other aspects, are designated intelligible "consequences" of the aspects used in the construction of the "ground." It will be seen that the classical criterion: "With the given ground, the consequence cannot be other than it is," follows with necessity from our conception of the intelligible relation of ground and consequence.

[84] The term "to read off" is meant to express that here a "seeing," a "gathering" of something from a thing is involved. Insight is in fact a seeing, a becoming evident of something. (See Chap. VI, specifically p. 82.) Evidence, then, is the objective aspect of what in a more subjective sense is called insight.

The examples so far were mostly of situational "foundations" defined relatively; i.e., the elements of the situation were defined not absolutely, but in relation to one another. E.g., by $a > b$, no absolute value or order of magnitude is prescribed for the a, but a region of magnitude merely in relation to b. Analogously, by the situation: "two straight lines intersecting twice," the positions of the straight lines are defined not absolutely, but relatively to each other. In contrast to these cases, a more primitive form of evidence or insight is involved if the situation is defined absolutely. Thus from the premises $a = 6$ cm. (or, say, 4-7 cm.) and $b = 2$ cm. (or, say, 1-3 cm.) it may be read off that "$a > b$". Or from the appropriately selected locations x_1, y_1 and z_1, and x_2, y_2 and z_2 of two objects a and b, it may be read off that a lies to the left of b, just as from two appropriate hues their 'similarity' may follow, and so on.

This form of intelligibility of phenomena which are based on given foundations has often been emphasized.[85] It is of the type of synthetic intelligibility; for the explicated relation need not have been realized when the situation was first perceived or conceptually characterized.

6. INTERPRETATION OF SYNTHETIC EVIDENCE FROM THE VIEWPOINT OF GESTALT THEORY

In a synthetic explication, a change of aspect occurs along with retention of—yes, of what, really? What is really to be understood by the "foundations" of an explication? In this connection, let us consider the case represented in Fig. 17. We are given a square and at a certain distance from it a line L, which one can consider as produced by the displacement of one side of the square so that it remains parallel to its direction. Here it may be read off that, if a point placed this side of the line L has the perpendicular distance α from L, it lies outside the

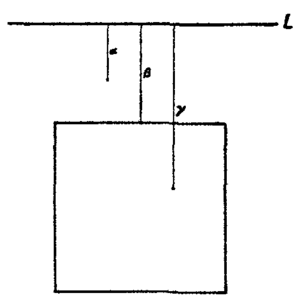


FIG. 17

square; that if it has analogously the distance β, it lies on the side of the square; that finally, if it has the distance γ, it lies within the square. Within each of these three statements, a change of aspect of the point takes place. Each time, the point is first characterized only by a definite distance from the line L. Then, suddenly, it is characterized by its position in relation to the square. Therefore, by virtue of a new "point of view", it has now altered its aspect. In terms of Gestalt theory: *it has altered its concrete "function" by the fact that it has entered a new organization.*[86]

Now, one might be tempted to general-

[85] This has been done, for instance, by Lindworsky (25, p. 37 ff). . . . Unfortunately, a genetic (temporal) priority of the founding to the founded contents has often been foisted upon their relation. Moreover, it was assumed, for instance by the Graz school, that these two kinds of contents have a different provenience. Thus, what was correct in their view of this relation was badly obscured. Often a relation is psychologically realized without awareness of its terms singly and *per se*. Such relation-experiences have this primacy over against their founding contents in common with Gestalt-experiences proper.

[86] One calls "function" or "part-property" that characteristic which appertains to an element only as a part of its concrete Gestalt-whole.

ize from this and to say: Synthetic evidence obviously consists in the fact that the new characteristics resulting from a new structuring of given elements are read off from these elements. But with this one would at once have lost again the specific meaning of evidence and of "insight." For if one properly combines pure hydrogen and pure oxygen (thus bringing them into a proper Gestalt-nexus), and if one reads off the result: the explosion and the formation of water, then to be sure, something very important is read off, but certainly no synthetic evidence. Or suppose that, in an experiment on perception, the brightness b appears equal to a second brightnes $b + s$, and this again to a third brightness $b + 2s$. If now in the comparison of b and $b + 2s$ inequality becomes apparent, something important is once more read off, but it contains no synthetic evidence. On the contrary, at first glance one is puzzled at such 'lack of logic,' even though one's own nervous system is here responsible.

In short, synthetic insight cannot consist simply in the fact that the result of a new organization is read off. In order to be comprehensible as an instance of synthetic evidence, the example of Fig. 17 must therefore be otherwise interpreted. Actually, with the change of its "function," the point remains not only numerically identical—so do the hydrogen atoms and the brightness b, as well—but it remains identical also as to spatial location. *But this means that it remains identical in the respect in which it serves as "foundation" for the two successive functions.* The position in which the point had been put by its function in the first structuring (it had the distance α from L) remains unchanged when in the second structuring the point assumes the function: "lying outside the square." We may express this as follows: The first structuring (the distance-relation to L) constitutes the foundation from which, in combination with further factors—such as the definite position and size of the square—the new aspect or new function is synthetically read off. "Foundation" therefore means a characteristic, constituted by the first structuring, which *must* be identically retained, if the new function is to appear evident in the new structuring. This foundation is what remains identical in the change of functions.—Incidentally, this identity brings about such characteristic phenomena as "following from" or "inner necessity." In reading off, one experiences how what is read off is founded exclusively on those properties of the situation which are constituted by the premises, and how it thus "follows from" these properties.

If this condition of identity be given sufficient attention, there will be no danger that the subject regard something as intelligible just because, in the moment of inspection, it happens to be apparent. Suppose, in the moment of reading off from: $a > b$ and $b > c$, the a were suddenly and maliciously to shrink (or turn red). Surely, under these circumstances the subject would not be doomed unsuspectingly to regard as evident from the premises that a must be less than c (or redder than c). For, foundations must enter into the new structuring with just those characteristics which they exhibit whenever considered in their earlier structuring.

Let us now summarize what has been attained in a final answer to the question of how synthetic evidence is possible: *Synthetic evidence is possible through the fact that from a situation given in a certain structuring and characterized by certain functions (aspects), without any* *change in essential foundations new functions (aspects) may be read off by virtue of new organizations.* By new functions, I mean functions which have not been utilized in characterizing the original situation.

With this, Kant's general question: "How are synthetic judgments *a priori* possible?" is answered, provided that "synthetic" is understood in the sense of following from non-constitutive co-containedness, "analytic" in the sense of explication through constitutive co-containedness, and *"a priori"* in the sense of intelligible.

Our answer to Kant's general problem is basically different from that of Kant himself, in that we do not reduce the synthetic *a priori* to prescriptions of reason incorporated into the object, but conceive this *a priori* as intrinsic in the nature of the objects themselves. Incidentally, concerning this point we are in agreement with Husserl's phenomenology.[37]

7. ON THE GENERALITY AND RELIABILITY OF SYNTHETIC EVIDENCE

This theory of evidence sets us among others this question: How is the generality of a given insight possible, although each time such evidence results from only *one* paradigm? For example, from a constellation consisting of a house, a tree, and a bush, how can I read off the general transitiveness of the relation "greater than?" From what source do I know that the result would have been the same if, instead of these objects, I had used as model a vase, an inkwell, and an eraser?

For this generality of a given evidence, obviously no more is necessary than that, in employing a concept—e.g., the concept "greater than"—I am able to distinguish the essential from the non-essential characteristics of an object. Thus I must distinguish the characteristic: "greater than" from the characteristic: "about 2.5 meters greater than," more particularly from the entirely non-essential greenness of the tree which is greater than the bush. I can in fact do this, as truly as I can conceive and employ concepts at all. Consequently, I can also recognize, during inspection, what characteristics of the foundation enter into the given evidence present and what others do not.

We have not yet clarified the relationship between synthetic evidence and the "inaccuracy of all inspection" (Schlick 35, p. 27 ff.). Why, for instance, is one unable directly to read off the fact that the sum of the angles of a given triangle is 180°? Why, in so many cases where exact magnitudes or coincidences are involved, does inspection generally permit only approximate conjectures?[38] In regard to certain tasks, inspection does actually prove to be too inexact. But if it is probably a correct statement "that all inspection or other experience lack complete clarity and exactness," one cannot conclude with Schlick that quite generally inspection lacks *sufficient* accuracy. For, one must first inquire whether, if clarity is not complete, it is also always too incomplete for a given purpose.[39]

[37] It is irrelevant here that Husserl divides our synthetic *a priori* yet again, and names "synthetic" only that knowledge which is founded in "material" essence, while he calls "analytic" the knowledge founded on "formal" essence. See "Logische Untersuchungen" II, 1, p. 251 ff; II, 2, p. 189 ff.

[38] To be sure, inspection has a heuristic function inasmuch as it offers suggestions. This function is of the greatest significance from the viewpoint of the psychology of thinking. (See Chapter VIII, p. 108 ff.)

[39] If the phenomenal data were too inexact in every respect, then postulational theory would also be futile. For here, too, actual performance consists in mental events.

Let us consider the usual proof of the theorem on the sum of the angles. Let us take it as proven that (see Fig. 18) $\alpha' = \alpha$, $\beta' = \beta$, $\gamma = \gamma$.[90] Now, *inter alia,* we still need the proposition: equals added to equals result in equals. Is a special clarity of inspection necessary for this proposition, i.e., for the reading off of the conclusion from the premises? Did we need some particular accuracy of inspection to observe that two intersecting straight lines do not come together again?—or that from $a > b > c$ it follows that $a > c$? (See especially the examples discussed in 9 below.)

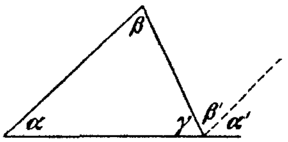


FIG. 18

There exist conditions, therefore, under which what is to be read off becomes sufficiently evident in the given situation. Only where this is not the case, does a naive person first look around for proofs. The original function of proof is to resolve something not directly evident into something which *is* directly evident or into a chain of directly evident steps.

An instructive example for this: On a mountain trip, on which descent was by the same path as had been the ascent of the previous day, I asked myself whether there must be a spot en route at which I must find myself at exactly the same time on the descent as on the ascent. It was of course assumed that ascent and descent took place at about the same time of day, say from five to twelve o'clock.—But without further probing, I could arrive at no conclusive insight. Since then, I have put this question to dozens of naive persons as well as of intellectuals, and have observed with great satisfaction that others had the same difficulty. Let the reader himself ponder a bit.—Certainly there exist several approaches to an evident solution. But probably none is, I might almost say, more drastically, evident than the following view of the situation: let ascent and descent be divided between *two* persons on the *same* day. They must *meet.* Ergo . . . With this, from an unclear, dim condition not easily surveyable, the situation has suddenly been brought into full daylight. The answer becomes evident, inspection sufficiently exact.

[90] In this proof it can be seen nicely how only those characteristics of the paradigm matter in the reading off which make the angles "angles on parallels" or the whole figure just a "triangle".

8. ON SYNTHETIC EVIDENCE FROM GIVEN, NOT FROM CONSTRUCTED, FOUNDATIONS

In the treatment of evidence so far we have dealt almost exclusively with foundations of evidence which were paradigmatically constructed from certain conceptual premises; in short, we discussed *constructed foundations.* However, it is clear that evidence is not restricted to constructed situations. Apart from the newly introduced "point of view" (p. 52), the foundations may also be simply 'given', 'found as existing'. They need only to be "characterized" somehow. After all, what matters in synthetic evidence is only this: the original way in which the foundations, their definite "characterization", are given in a situation, should not already constitutively cocontain what is to be read off. What is read off must represent an aspect which is new in contrast to the original phenomenal aspect—not of course new in the sense of added material such as may be discovered by more exact observation.

Instead of further discussion let us have an example: Given a particular room with its furniture, etc. Here the following statement may hold, for instance: "If, in this room, one goes directly from the stove to the door, one passes close by a chest of drawers." This aspect need in no way be constitutively co-contained in the original way the room was given (as was, e.g., the presence of the chest of drawers); but it may nevertheless be read off from the facts comprising this room with the straight line drawn therein.—Another example, which involves an important component of almost all practical manipulations.[91] There is given an object at a definite location P. Now if my body, my hand or a stick which I am wielding moves in a definite direction—not characterized with reference to P—it may be read off that my body (or hand or stick) will come into contact with the object at P, provided that nothing as yet unforeseen interferes.

The sole difference between these two cases and our earlier examples is that now the inspected situation is not constructed from any concepts, but for the most part is simply 'given'.[92] This holds in the first example except for the straight line and the consideration of its surroundings, in the second except for the movement in the particular direction. The reason for the emphasis in this chapter on the special case of reading off from a *conceptually constructed* foundation is simply: What is peculiar to synthetic evidence is more easily demonstrated if the foundation is conceptually 'in hand' to begin with. About the pure state of being given, no communication is possible, and the "how" of such a state allows again of only conceptual characterization.—It contributed to this emphasis that I wished to facilitate the connection with older treatments of the problem.

When we are no longer limited to conceptually constructed (paradigmatic) foundations, the field of application of our concept of evidence increases tremendously. In daily life, we do not read off from explicitly established premises, but from given situations and from operations therein undertaken. In practical life, we are not interested in general statements which allow of general formulation and general understanding, but in the fact that something holds here and now. (Not to mention that, by the requirements of conceptually constructed foundations, animals would be excluded at the outset from the province of evidence, which would be unfair.)

9. EXAMPLES OF SYNTHETIC EVIDENCE

In the following, a few further examples of synthetic evidence will be added, in order to demonstrate the range of application, the implications of the theory.

The first three examples involve not only a single reading off, as was the case till now, but a whole chain of evidences which terminates in a "final reading." In such a chain, the results of the earlier evidences enter into the construction of the foundations of the later ones, up to the final reading which gives the conclusion.

1. "An odd number plus an odd num-

[91] It appears, for example, in all solutions of problem situations of the type which Köhler investigated.

[92] This does not mean that the foundations may not also be constructed, and indeed heterogeneously so, i.e., without utilization of the relation which is later read off. For this purpose, one need only characterize the positions of the different components of the room by means of their spatial coordinates in some reference system. And even though the translation of these various coordinates into the corresponding visual distribution within the room would constitute a relatively tedious process, this translation would still be of the same type as the translation of the expression $a > b$ into a perceptual model.

ber equal an even number." "Odd signifies a quantity of pairs plus a half pair." In the first place, it may be read off that two quantities of pairs, combined, result again in a quantity of pairs. It may further be read off that a half pair together with another half pair results in a whole pair. Finally a third reading leads to the conclusion.[93]

2. "$7 + 5 = 12$." The premise "$7 + 5$" constructs the joint quantity (at first consisting of x elements) from a set of 7 and a set of 5 (other) elements. The question, "How many is that?" (What number is x?) requires a definite operation, namely a counting off of this joint quantity so that each element is named exactly once, and the naming of the number which falls on the last element. From the joint quantity so counted off, more exactly from the last element of this counting, the name "12" may be read off. (Let it be noted that here we are not concerned with definitions as they are used in arithmetic and in the theory of sets. We are occupied solely with a description of the thought-process by which such a statement as "$7 + 5 = 12$" becomes evident.)

In the next two examples, physical situations are treated.

3. If a mirror is rotated by an angle δ, and if the angle of reflection always equals the angle of incidence, then the reflected beam rotates by 2δ. It may be read off that, firstly, the rotation of the mirror displaces the old angle of reflection as a whole by δ and therefore the reflected ray by the same amount; and that, secondly, the angle of incidence, enlarged or diminished by δ through the mirror's rotation, enlarges or diminishes the angle of reflection—as its equal—again by δ, so that (final reading) the reflected beam rotates by 2δ.

4. Greater volume with constant mass means less density. From the expansion of a constant mass (e.g., of a gas), it may be read off as a "new aspect" of the situation that the substance concerned has become "thinner" throughout.

In conclusion, one more example, from pure logic:

5. The forms of logical conclusion as well owe their intelligibility to synthetic evidence. For example: From "all M are P" and "all S are M" (or, "this S is M"), it follows that "all S are P" (or, "this S is P").—Here the conclusion arises by means of a different relating of the terms; compare the representation by areas, after Euler, in Fig. 19.

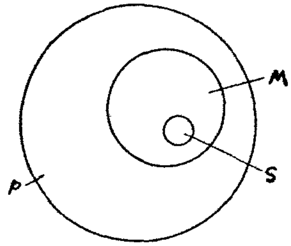


FIG. 19

Incidentally, on the intelligibility of this procedure depends the intelligibility of the application of every general rule. I write the "formula of application" as follows: "From $f(a)$ and $a = b$ follows $f(b)$." E.g., to take (f) the longest object (a); the longest object is this wire (b); ergo: to take the wire [$f(b)$].—If a human being or an animal, say, in a training experiment, has grasped the principle of the required conduct, e.g., "always to go through the door next to the middle," he will from now on act with insight in regard to that principle. Every intelligibility dependent on nothing but intuitive or logical deducibility from a general law (see p. 5), is exactly of the type of

[93] The new aspect arises here each time through the fact that two parts are inspected "as a whole" (see p. 52 under d).

evidence involved in the above procedure of application.

10. APPLICATION TO THE PSYCHOLOGY OF THE UNDERSTANDING AND THE FINDING OF SOLUTIONS

This theory of total insight has purposely been developed without reference to the experimental material. It had to be mature in itself, before it was applied to the set of problems with which we are specifically concerned.

Let us begin with the most simple. In so far as analyses of goal and of situation consist in explications from the proposition and the premises, they are of the type of synthetic or analytic reading off.[94]

We have seen that the "point of view" in reading off (see pp. 52-53) is each time prescribed by the *Problemstellung* as a whole. Suppose that from the proposition that the prime numbers constitute an infinite series, it is read off that for every prime number there is always a greater one. In this the "point of view" of the reading is dictated by the wish to transform the proposition of infinity, which is not accessible to a finite method of proof, into a finite and general proposition. The realization that it is the infinity-factor which makes the original proposition incapable of proof is reached through the analysis of conflict. This analysis presupposes merely an awareness of what is really meant by proving something. And this again means a most general analysis of goal (of the type of analytic insight).[95]

The rôle which such very general and purely analytic goal-explications play in problem-solving is as decisive as it is hidden. In the problem of prime numbers, the general nature of proof was to be analysed. In the following problem (discussed by M. Wertheimer) it is the general nature of measuring. Task: To express (measure) the side of the square in

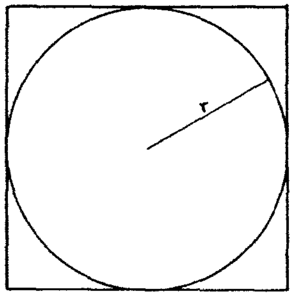


FIG. 20

terms of the radius r of the inscribed circle. (See Fig. 20.) Ideal solution: To measure means to make the measure coincide with what is to be measured. In the present case, the nearest approach to coincidence is parallelism and therefore rotation of the radius into the position parallel to the side. One sees now that the side of the square equals $2r$.

Through the special choice of the examples thus far used, the impression could easily have been given that synthetic evidence was pretty well restricted to the province of pure mathematics and logic. However, it is sufficient to bring to memory that so-called pure mathematics and logic is for the most part (i.e., with exception of the "artificial" axiom-systems and the theorems derived therefrom) embodied in reality; in other words, that everyday life is full of mathematical and logical situations. Let us take the two practical problems which occupied us at the very beginning, the radiation problem and the pendulum

[94] The expression, "analysis" of situation or goal, ought not to be confused with "analytic" reading off. Synthetic reading off, too, represents, as has been stated, an "analysis" in the broader sense of the word.

[95] We named a reading off "analytic," in the narrower sense, if it does not "affix" a completely new aspect to a given constitutive one, but merely explicates what inheres in this given aspect.

problem. In both cases, the understanding of the best solution involves synthetic insight at the decisive point. In the first case, it is decisive that more material exists at the intersection of several streams of material[96] than elsewhere. This accumulation of material may with perfect insight be read off from the model of the intersecting streams. The almost startling evidence of the solution by crossing depends on this "synthetic reading off."—

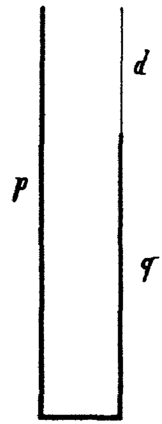


FIG. 21

The best solution of the pendulum problem owes its evidence to the following synthetic reading off: "If p and q (in Fig. 21) expand by the same length, then the distance d between their end-points remains constant."—Let the more material aspects of the effects, namely, that the rays destroy organic tissue, and that warming causes expansion, be as purely empirical and unintelligible as they please, the more "formal" aspect follows both times with entire intelligibility from the geometry of the rays or that of the pendulum, as contained in the solution.

It is generally a valid statement that facts (aspects) can be *intelligible* on the ground of synthetic reading off only in regard to their wholes. It is intelligible that the stick moved in this direction will strike the ball lying there, if the total movement of the stick, including the future part, be conceived as a non-temporal line, and if this line be related to the position of the ball. But what the ball will do when struck is by no means totally intelligible from the premises just named. The physical reaction of systems to one another contains, as we shall see (see Chapter V), certain *components* which are "merely to be accepted", in other words, which are not intelligible.

We have recognized that the intelligibility of the best solution of the radiation problem, as of the pendulum problem, depends on synthetic evidence within the given situation. Therefore, we are dealing here with a much more profound understanding of the solution than was that more general form of understanding discussed in Chapter I, 6 (see also p. 63). There it was in principle sufficient for understanding if, in the form of "functional value," connection with a sufficiently simple and general causal relation between solution and goal had been established (see the oxygen example, page 5). This underlying causal relation in itself could be totally unintelligible.

Even the most general heuristic methods of productive thinking are of the type of evident explications, more particularly, of the analytic kind. Thus, for example, analysis of situation, where it consists in deduction from the given premises, itself derives analytically from the nature of proof. For, "to prove" really means nothing but deduction of what is asserted from the given premises. Still more generally: Both analysis of situation as examination of what is given,

[96] For the sake of simplicity, a bundle of rays is here considered as a stream of material, i.e., by analogy with a river.

and analysis of goal as examination of what is demanded, these two fundamental methods of rational discovery, themselves derive analytically from the nature of problem-solving in general. For, to solve a problem involves making what is given serviceable to what is demanded.[97] That I must know what is given, in order to operate with it, and must also know what is demanded, in order to operate toward it, this in turn follows analytically from the nature of action. For action means acting, guided by knowledge of the purpose and of the means. This structure of action is no "definition," but an original and basic experience of mankind. Also the question, "Why doesn't it work?" or, "What should I alter to make it work?," in other words, what we call analysis of conflict, is evident—evident from the nature of "solution." For solution is ground of the realized goal, thus elimination of a possible obstacle to the goal.

But an important question yet remains to be answered. Assumed that the goal is intelligible from the solution, i.e., that the solution is entirely comprehensible as a solution. For this reason alone, could the solution already be found with complete insight? No, because for this, the solution in turn would have to be intelligible from the goal. After all, it is the goal which is given; the solution is yet to be found. But to the intelligibility in the direction from a to b, there corresponds in general no such evidence in the reverse direction. (Thus, for example, from $a > b > c$, $a > c$ may indeed be read off, but not the reverse: from $a > c$ it does not follow that $a > b$ or $b > c$.) Only in certain limiting cases, where goal and solution are equivalent to each other, does deducibility exist in both directions (cf. analysis of goal in the sense of a "reversible" deduction from the proposition given). But these are exceptions, which—be they ever so important in their place—do not affect the rule that in most cases synthetic evidence has for thinking the character of a one-way street, which is not open to traffic in just the crucial direction: goal to solution.[98]—But does the intelligibility of a solution have thus no heuristic significance at all?

Such a conclusion would be overhasty. How about the approach to a solution *via* the pertinent ground of conflict? (See Chapter II, 3.) Without doubt, wherever from a solution a the goal b follows intelligibly, there the conflict non-b can be read off from the corresponding conflict element non-a. A conflict element, however, differs from the corresponding solution by the extremely important fact that, along with the conflict itself, it is *present in the given problem-situation,* and that it is therefore in principle accessible through analysis of the situation. Now the decisive question arises: Is a conflict element more easily discovered if it bears *intelligible relation to the conflict?*

Let us consider the situation of the stick-door problem (p. 26). Without doubt, from the conflict *elements* in this situation—from the oblique position and the length of the stick, from the narrowness of the door, etc.—the conflict finally to be eliminated, the crossing, follows intelligibly. *But now, in turn, could some*

[97] We may leave undecided whether this is actually "analytic" and not rather "synthetic" evidence. For, this depends on whether or not the situation in question was introduced (constituted) by another aspect (see p. 48-49); and this is to a certain degree a question of the particular psychological case.

[98] In comparison: one can understand and feel quite deeply a poem, a song, or a picture—without oneself being able to produce such things.

indication of the conflict elements be gathered from the given fact of the crossing when the former are sought?

The question thus modestly formulated must in any case be answered affirmatively. For no one will wish to deny that the fact of the crossing characteristically singles out its essential foundations—just the position and length of the stick, as well as the narrowness of the door—in contrast to such irrelevant properties as, e.g., the colors of stick and door, and the like. In other words, the crossing as a spatial relation singles out its spatial relata before other factors, e.g., color factors, just as a color-relation would single out its color relata before other, e.g., spatial factors of the situation.

I am not yet completely clear about the nature of such singling out, which undoubtedly exists. It does appear to be again of the type of insight due to reading off. The important question, whether intelligible solutions—i.e., those from which the goal may be read off—can in principle also be found with full insight, is to be answered affirmatively in any event. (In Chapter II, 7, we have already become acquainted with certain factors which decide whether, and how easily, thinking hits upon just one particular factor among the many conflict elements present. What was said there is of course valid for intelligible conflict elements as well.)

In the following chapter we shall, among other things, investigate this problem: There are situations which do not allow of total insight in the sense of synthetic or analytic reading off—as, for instance, in the province of causation in a strict meaning of the word. Is at least some degree of insight possible in such situations? And if so, what is the nature of insight under these circumstances?

The intelligibility of mental events through so-called empathy may be mentioned here only in passing, as it is of little consequence for our subject of problem-solving. Clearly, this much-discussed intelligibility is connected with the fact that causation may be experienced in one's own mental processes. Here and only here is there coincidence of causal relations and the experiences of a subject. Good examples of this "inner" intelligibility will be found in W. Köhler's "Gestalt Psychology", Chapter X. (21)—About another instance of evident following-from, see footnote 76, p. 49 of the present book.

Chapter V

On Learning and Partial Insight

1. Empirical Connection-Structures

According to the preceding chapter the intelligibility of many connections of the type: "if a, then b" rests on the fact that b can be read off when a is given. This evidence is either synthetic, i.e., b is a new aspect of the heterogeneously constructed situation a, or it is analytic, i.e., b is a constitutive aspect of the situation a, and therefore not "new." —Now questions like these arise: *Of what kind are those other if-then connections in which the "then* b" *can in no way be read off from the* a*? How are such connections accessible to a thinking creature? In particular: Does all intelligibility, in the sense of the general definition in 2 of the preceding chapter, depend on the possibility of such reading off?* Suppose that in a given case this possibility does not exist. Then the following other possibility must still be examined: Among all the phenomenal characteristics which b could assume, those which b really has may be directly singled out by the phenomenal nature of a. The term "directly" is meant to exclude the intervention of further factors.

Let us begin from below, i.e., from that minimum of "rationality" which functional connections in a world must exhibit if a creature is to be able to penetrate into this world by thinking, and thus to get along in it.

If the spatial and temporal connections among the events of the world were without any rules, were completely chaotic and thus absolutely a matter of chance, then thinking would have no practical significance at all. *Constancy* of connections is thus the minimum of rationality.[99] But since no event repeats itself quite identically, a constancy of connection can prevail only among parts or aspects of events. For only such repeat themselves. For thinking, then, the problem arises of discovering what components of connected events have a constant connection, or—phrased somewhat more familiarly—what properties of given data are causally essential for an effect on which our eye is focussed.

Now in our actual world, such constant connections (laws) have for the most part not the type of a constant coupling of definite "elements," but that of a *constant structure of variables.* For example, there is no absolutely constant connection between the boiling of water and a quite definite temperature. But there is such a connection between the boiling of any fluid and that temperature at which the vapor-pressure of the liquid just exceeds the atmospheric pressure at the place concerned. In other words, the elements, such as the liquid, the temperature, the place, the pressures concerned, can vary, and yet the constancy of the connection is not thereby affected. On the contrary: Under altered conditions, events must occur which are altered in a definite way, so that the same connection-structure may be retained. (Compare the transposability of a gestalt, especially the transposability of a comprehended solution, as discussed on page 5.) Not only in science, however, but everywhere in common life, man (and in principle, the same holds for the animal)

[99] The statistical conception of natural laws does not deny constancy in general, but only constancy on the microscopic level.

finds himself confronted with the problem of grasping the constant structure of interrelated variables, *of "getting the idea."* For example, a chimpanzee will again and again promptly and sensibly utilize appropriate stick-like tools (e.g., a piece of wire, a wisp of straw, a shoe) when he wants to draw near a goal-object which is out of reach. He is capable of this only because he has grasped that the point is not the color, the original position or the absolute spatial direction of the tool, but purely and simply this structure: to move the goal within reach by means of a sufficiently long object. [100]

2. LEARNING OF CONNECTION-STRUCTURES

Now it is very important that the problem of grasping the essential can also be solved in cases in which the connection-structure as such lacks any intelligibility; in which, therefore, for the subject in question any other effect could be just as well connected with the given cause as the actual effect. Nature allows its creatures to grasp all kinds of constant connections into whose inner necessity she still forbids them to penetrate. Among the philosophers, at least since Hume, there prevails the almost unanimous conviction that in physical nature none but such entirely unintelligible causal connections will ever be found.

We first raise the question: *How is it possible to grasp connections which are constant but in themselves totally unintelligible?* This question has already been answered by Francis Bacon and John Stuart Mill in the general form of so-called rules of induction. Their quintessence may be expressed as follows: *The essential cause of an effect b—or that which "leads to b"—can be grasped by the abstraction of those further factors which all b-situations have in common, and of those which all comparable non-b-situations lack in common.*[101] Such an "abstracting induction" takes place everywhere in the practical life of man and animal, only less systematically than in the sciences. Psychologically it means a process in which, out of a number of situations of a class, *the common aspect is "precipitated."* The result of such a "precipitation" is a *change of aspect* of the situations involved. This often occurs quite suddenly, possibly accompanied by an "Aha-experience."

Reorganizations of this sort may be very nicely observed in training experiments on man and animal. One can, for example, train chickens always to select their food from the lighter of two grey papers, whatever the absolute greys. Or one can train a person always to choose the box next on the left from the middle one, with varying position and number of the boxes offered. Such training often leads to sudden "insight." The principle of the required behavior is all at once clear, and the curve of errors exhibits that famous and much-discussed *sudden drop*. Yet the principle *per se* may be as unintelligible as one pleases.[102] There-

[100] A few fine examples of use of a stick *without* understanding of this functional structure have been given by K. Gottschaldt (9, p. 124 ff.).

[101] Huang's experiments (13) bear eloquent witness to the degree to which this criterion is autochthonous. Huang, in experiments with children, brought about opposite effects by apparently the same event. For example, a needle laid on water sank once and did not sink another time. It is amazing how the most harmless differences between the two situations were promptly seized upon by the children as "cause" for the varying behavior of the needle.

[102] Among actual causal connections, those with which the experimentalist in medicine must so often be contented most nearly approximate the extreme type of totally unintelligible connections. (See in regard to this de Kruif's excellent

fore, paradoxically expressed, there may be insight in a connection which is totally inaccessible to insight. The paradox is resolved as soon as one realizes that here the term "insight" appears in two different references. By the grasping of the common principle of a number of situations, this principle—its inner Why—is not comprehended. What is comprehended is that once this common principle is given, *the individual situations must be just as they are and not otherwise.* The different situations, till then seemingly accidental, suddenly appear throughout as embodiments of a single principle. Therewith they acquire that "intelligibility" which has already been discussed in Chapter I (p. 5), and which consists in nothing but *"reducibilty to a general law."* (See also Chapter IV, 9, example 5.) This may be called "insight of the second degree."

Unfortunately, through this ambiguity of the term "insight," great confusion has arisen among psychologists. One speaks of the one meaning when he thinks he is speaking of the other. For the sake of unambiguity, I designate a connection of the type: "if *a*, then *b*" as intelligible, *not* if it is grasped as a principle common to a number of data, from which the latter can be derived, but if and in the degree to which, before other possibilities, *b* as such is directly favored or singled out by the *a* (see Chapter IV, 2). This will be called "insight of the first degree."

3. VERY GENERAL CONNECTION-STRUCTURES

It is obvious that, with the help of abstracting induction, a principle will be the more surely and quickly grasped, the greater is the number of situations from which the identical principle may be abstracted. Now in this respect, the constitution of our world is relatively convenient. Let us consider the situation cited above, in which some stick-like object was utilized for drawing near a distant goal-object. This behavior presupposes, among other things, the general experience that one thing is *movable* by another thing. Plainly, in order to have this experience, a subject does not need just the stick-goal situation and all sorts of trial-and-error reactions therein. Some other instance of communicated motion, not necessarily one related to a goal but perhaps a playful one, could demonstrate exactly the same to him. It is also unnecessary that it be the learning subject himself who directly or indirectly moves something; rather, transmissions of motion occur continually in our environment, as things push and pull each other. From this, time and again, the same lesson is to be drawn. To be sure, there are "objects" which are able to interpenetrate unhindered, as for example waves, shadows, and the like; but they are after all notorious exceptions.

Still much more general is the experience—realized, by the way, in both the radiation and the pendulum problem—that an effect depends also on the nature of the reagent (on its "sensitivity"). Even the infant, kicking and tossing about in his crib, experiences that it makes a difference whether one strikes the cover or the wooden railing. And it does not remain hidden to him that the mother reacts to crying differently from other people.

In summary: Everywhere in our world the particular causal relations participate in very general ones. But "very general"

book, *Microbe Hunters,* which is very rich in information about thought-processes, for instance, about the heuristic significance of chance.)

means: accessible in nearly every occurrence. Thus abstracting induction is surely not starved for material.

4. ON PHENOMENAL CAUSATION: A) COINCIDENCE IN SPACE AND TIME

This much is beyond doubt: even in a world of totally unintelligible connections, a subject can learn what matters in each case and can solve new problems on the basis of such general experiences. But, by definition, what a subject cannot do in a world of totally unintelligible connections is this: from the observation of an event he can derive no immediate indication of the nature of another event which is causally connected with the first. If we want to know whether the causal connections of our world are —apart from really evident aspects (Chapter IV, 10)—at least *partially* intelligible, we must therefore ask the following question: *Does, for instance, a demanded effect contain something which in any way singles out the (requisite) cause from other conceivable events?* If and in the degree to which this holds, two things would follow:

1. Even in the province of natural causation, thinking would not be absolutely blind in its heuristic efforts.

2. The learning of causal relations would be facilitated through the contents of the things to be connected.

We shall take the stick example as basis for our discussion. The goal was here to get hold of the banana; the solution: to hit the banana with a stick moved in the proper direction. Obviously, if one wishes to have a banana, one must in any case do something about the banana itself, i.e., one does not strike, say, at a stone 2.7 meters to the left of the goal, but just *at the latter itself*. From this follows at the same time that it is the length and not, for instance, the color of the tool which matters.—Now this, like every case of "contiguous action,"[103] actually involves partial intelligibility; at least the place of the cause is singled out before other conceivable places by the place of the effect (just by virtue of the coincidence). The places of cause and of effect are not "accidental" in their relation to one another.[104]—Now since, due to its special simplicity and "prägnanz," the principle of "contiguous action" is an outstanding case among many principles conceivable in its place, it need not be so painfully learned as any entirely unintelligible connection. That sometimes no learning at all is needed, follows from certain primary reactions of all creatures, namely their "instinctive"[105] goal-directedness.

The *spatial* coincidence of cause and effect is only one side of the complete principle of "contiguous action". The other is that of *temporal* coincidence, more exactly: the coincidence of cause and effect as singularities, discontinuities in temporal developments.—Someone comes home of an evening. A gust of wind slams the door shut behind him. At the same moment at the other end of the corridor, the light goes on in a room whose door is ajar. Although one knew ever so well that no causal connection exists between the door's blowing shut,

---

[103] [*"Nahewirkung."*]

[104] Incidentally, this same spatial coincidence of cause and of effect, or of agent and reagent, is in many cases the ground for the fact that (see p. 20) the solution deals with the material of the given problem-situation, in the sense of altering just this material. At least in practical problems, the heuristic method of situation-analysis thus receives its decisive meaning.

[105] This word is placed in quotation marks for a good reason. For, with the superficial alternative: either product of learning or instinct—it is usually overlooked that facts of simplicity such as the coincidence of cause and effect may also be of consequence for the genesis of an instinct.

and the light's going on, that rather someone in that room has turned on the light, by chance at exactly the same moment—still he would be unable to escape the compelling impression of causal relation. So decisive is the temporal coincidence. *We generally perceive as "cause" of an event, of a singularity, another singularity which coincides spatially and above all temporally with the first. This in its turn results as "intersection" of two uniform developments or "world lines".*[106] Thus the going on or off of the electric light switch coincides with the intersection-point of the world-lines of the light switch and the moved arm; the street's growing wet coincides with the intersection-point of rain and street, etc. We have here before us the real *"gestalt factor of phenomenal causation".*[107]

In summary: *Phenomenal causation in our world owes to the law of contiguous action a remarkable simplicity (prägnanz) in spatial and temporal respect. At least in regard to their positions in time and space, cause and effect are related not at random, but intelligibly. The time and place of the cause coincide phenomenally with the time and place of the effect.*

[106] With Minkowski, one calls "world line" the four-dimensional, i.e., the spatial and temporal, extension of an object.

[107] The factor of *temporal* coincidence of cause and effect is in fact the temporal variety of a familiar gestalt factor in figural grouping (see Max Wertheimer's investigation, 43), according to which two neighboring discontinuities in an otherwise homogeneous field appear compellingly as a pair. To this end they need not resemble each other in *content.* They may even be more similar to the surrounding field than to each other. Their "similarity" *qua* "figures", i.e., their common prominence, their common standing out against the ground, suffices in conjunction with the "factor of proximity".

By the way, the factor of temporal coincidence is analogous to the criterion of causation discussed on page 64. There, however, it was a matter only of "coincidence" of *conditions* rather than of *events.*

5. ON PHENOMENAL CAUSATION: B) CORRESPONDENCE OF FORM AND MATERIAL

At least as important for man's dealing with causation as those spatial and temporal correspondences of *position* are certain correspondences of *form* between cause and effect or agent and reagent. Such correspondences as to form necessarily arise from constant temporal and spatial relations between cause and effect. An example of temporal correspondence of form: the rhythm of the sounds of knocking corresponds to the rhythm of the motions of knocking. An example of spatial correspondence of form: the trace resembles the object which leaves it. Analogous correspondence of form exists between agent and reagent: the wards of a key resemble the keyhole, the (shortest) detour resembles the contour of the obstacle. Cause and effect also correspond with regard to the form of change; i.e., the variations of cause and effect parallel each other. On the whole, increased destructive effect (see the radiation problem) corresponds to increased radiation-intensity; increased displacement or deformation corresponds to increased intensity of impact.[108]

The correspondences of form which we have just mentioned constitute a special case of correspondence as to "content" in general. The same stability of our world which is illustrated by the fact that the majority of objects in our environment remain relatively unaltered, manifests itself also in the observation that, many properties tend to pass unaltered from the cause into the effect. Such characteristics are preserved when an occurrence spreads from one system to another

[108] With this, a further intelligible aspect has been shown up in the material of the stick example.

system. Thus the wetness of the rain becomes the wetness of the street, and the color of the light becomes the color of the illumination. Again—see the stick example—the motion of that which pushes is continued in the motion of what is pushed. Generally speaking, not only the phenomenal genus of motion as such is here preserved but also to a degree its direction. Moreover, we have to consider what W. Schapp (33) has called "perceived causation": heavy things make 'heavy' noise, dainty things move daintily, etc. The "qualitative" causation of the primitive, magic world is so to speak an exaggeration of such instances of intelligibility; e.g., what is fast makes fast, what is strong makes strong, etc.

It will be seen that *cause and effect are related intelligibly not only in regard to "position" but also to a high degree in regard to "content": properties of form, character, direction, material, etc., pass right before our eyes from the cause into the effect.*

All these instances of intelligibility consist in certain relations of identity or similarity between cause and effect. Obviously, such instances of "partial" intelligibility in causation are of a nature quite different from that of the analytic and synthetic intelligibility on which depends what we called "total evidence" (see Chapter IV). Such partial intelligibility also bears not so much the character of necessity as of *simplicity, prägnanz.*

With this, the second of the questions set up at the beginning of this chapter (see 1) has also been answered. There are cases in which no complete intelligibility, no reading off, is present, and where nevertheless "*b*'s real phenomenal characteristics in contrast to other possible characteristics appear directly singled out by the phenomenal nature of *a*". Hence, intelligibility is not everywhere based on the possibility of "reading off". And even if the "partial" intelligibility with which we just became acquainted is far less pretentious than the "total" intelligibility based on reading off, it is by no means so unpretentious that it could not justly be contrasted with the statement of Hume quoted on page 48. Certain features of an effect may in fact be at least "conjectured" from the cause. And for the finding as well as for the learning of causal connections, this is extraordinarily important.

Incidentally, most such correspondences between cause and effect are to be expected only where the causation does not consist in pure "release" or in "concealed" occurrences. It became possible to overlook to such an extent the rule: *"causa aequat effectum"* because—after the famous examples of Bacon, Hume and Mill—thinking was too onesidedly concerned with cases of "release" or with processes like chemical reactions, essential parts of which are inaccessible to observation.[109]

6. PARTIAL UNINTELLIGIBILITY OF CAUSATION IN NATURE

It is obvious why the term "partial unintelligibility" has here been used. By no means all properties of the effect are phenomenally founded in attributes of the cause. Certain unintelligible aspects enter into every causal connection in nature. Thus, for example, in using a stick as a tool, the fundamental facts of movability and impenetrability of objects is something to be accepted as mere fact. If one were to object that the im-

[109] The contents of 4 and 5 constituted the main part of my report on *"Lernen und Einsicht im Dienst der Zielerreichung"* (4).

penetrability of the objects has already been suggested by their behavior at touching and handling, the reply would be: certainly, but this behavior itself is after all the same experience of impenetrability in another form. This experience itself does not allow of further rational reduction.

The category of "sensitivity", too, is only partially intelligible, namely, only insofar as the dependence of the effect on properties of the reagent seems more sensible than would be a dependence on the properties of any other, indifferent, object. In other words, only the factor of contiguity is here intelligible. Otherwise a world in which as a rule the effect depends only on the agent would be entirely conceivable.

The intelligibility of the formulation: "The effect depends on the 'sensitivity' of the reagent," rests of course on purely analytical explicability. For, "sensitivity" is defined as that in the reagent which co-determines the effect.—On such analytical explicability are based innumerable so-called *Wesenssätze,* e.g., that death belongs to the structure of life. If life is understood concretely in its pervasive reference to death, the saying is analytical. If life is characterized simply by certain external properties, death is a "stubborn fact." (See also the analytical examples on p. 60, as well as footnote 76, p. 49.)

All unintelligible features of causation can of course be mastered only through learning by abstracting induction. However: such learning is immensely facilitated by the intelligible factors which accompany causation everywhere. In a world, for example, in which the effect did not coincide with the cause but depended on it according to some other—constant or varying—time-space function, learning might be extremely difficult and unfruitful.

7. STICK EXPERIMENTS WITH INFANTS

The strange interlacing of intelligible and unintelligible factors in the use of a stick strongly invited an experimental study of the genesis of this procedure. We owe important psychological information on the use of a stick to several works on animal and child psychology (see for instance 20 and 46.) But what was still lacking was a thorough genetic study. In the summer of 1931, by courtesy of the nursery school of the Kreuzberg District (Berlin), I was enabled to realize my wish to study in detail the development of the procedure in human babies. The Ss were eight children from 8.5 to 13.5 months old, none of them yet able to handle a spoon when eating. The problem-situation was always the same: the child sat at a table on which—out of reach—lay some attractive object (a red ball, a celluloid frog, or the like) and—within reach—a stick. If during the experiment the goal-object lost its power of attraction, it was replaced or supplemented by others, or given to the S for a brief time.

Results

1. Only two of the eight children (10 and 13 months old) used the stick as a tool either at once or after they had assured themselves by vain attempts that they should not reach the object simply with their hands.[110]

2. To the other six children, the stick

[110] With the younger of these two children, I was able to ask the parents about possible previous stages of the performance. I learned that the child had much association with older siblings and already played at home quite well with hand broom, spoon, and the like. In addition, he had once been an interested observer of an experiment with another child.

was at first a *plaything*, with which they banged on the table or on the wall, or which they enjoyed throwing time and again on the floor. In between they reached for the object, but reaching and playing with the stick were at first two completely separate occupations.

3. With four of these six children, the play with the stick had, or acquired, reference to the object. The gaily colored thing on the homogeneous ground of the table was bound to attract their attention over and over again; and it is not surprising that under these circumstances the playful movements of the stick were directed toward the object rather than into empy space. With three of these children there developed a regular play-at-a-distance.

4. Naturally, with such play it practically had to happen that the object would occasionally come plainly nearer to the child, sometimes near enough for him to reach it. With two of the four children (mentioned under 3) such a chance event induced abrupt recentering: *the stick-toy became a tool in the service of the striving for possession.* Formerly, reaching for the object and playing with the stick (and so with the object) were two separate concerns; now the play at a distance was completely supplanted by possessive desire. The wielding of the stick became a supplementary part of reaching.—With the two other children, however, no such chance event succeeded in "instrumentalizing" the stick.

5. The last two children (of the six mentioned under 2) played with the stick without any reference to the object. But with one of them an accidental contact which he noticed achieved at one stroke—i.e., without mediating play-at-a-distance—the instrumentalizing of the stick. The other child (8.5 months), in spite of accidental contact (and even one instance of accidental success), remained completely centered on his play with the stick, whenever he was not actually reaching for the object. For this child, by far the youngest, the gulf beween object and stick was deepest.

6. Four weeks later, when I wished to repeat and to film the experiment with this last child, to my very great astonishment I had to recognize that use of the stick as a tool, of which I had seen no indication the first time, was now perfect. I should not like to explain this solely from the "maturation" which took place in the meantime. Those chance effects had probably not occurred without leaving any trace.

In these findings, the following seems to me especially important: *As a rule, use of the stick as a tool arises from playful commerce with the stick and—by way of the stick—with the object. An approach of the object which accidentally occurs in this situation leads to the discovery that the stick-toy is suitable as a tool.*[111]

From Yerkes' monograph, *The Mind of a Gorilla* (46), I take a development of this procedure which is quite different in kind, much clumsier, but also observed under very different conditions. When all relevant findings are ordered chronologically, the following picture results:

January 5-9:[112] Stick indifferent. (The stick lay in the vicinity of the goal.)

January 9-13: Stick already interesting.

January 15-19: Only if the stick is in accidental contact with the goal does the solu-

[111] The fact that two of the babies used the stick as a tool at once somewhat loses importance when one remembers that with the younger of them, lively handling of stick-objects had already been observed earlier, while those three of the six children whose parents I could question had not yet been seen to use such objects.

[112] From Jan. 7, problems were given along with this in which the goal was fastened to a string.

tion follow (and then in the stereotyped form of sweeping movements of the stick toward the left).

January 19-26:[113] The nearby stick is spontaneously brought into contact with the goal. The automatism of the sweeping movement is replaced by varied manipulations of the stick.

January 27-29:[114] The solution follows only if the stick already protrudes through the bars of the cage. (In the beginning, it was even necessary that most of the stick be protruding.) On January 29, the stick is by chance pushed through the bars, and this immediately leads to the solution.

January 30-February 1: From now on the stick is spontaneously pushed through the bars. Even a stick found somewhere in the cage is utilized.

February 2-4:[115] Sticks are looked for. Even other objects (chain, straw . . .) appear as substitutes.

(About fifteen minutes were spent on the stick problem each time.)

8. SOLUTIONS IN CONNECTION WITH THE "INSTINCTIVE" REPERTOIRE

A few more examples will now illustrate the importance of varied commerce with things and situations for problem-solving. They will also show how the kind and range of such commerce are determined by the constitution and the way of life of a species.

It is certainly not without importance for the genesis of stick-solutions by chimpanzees when Köhler reports: "Straws and twigs are also used as spoons in pure 'play' . . ." (20, p. 75). "A 'halfway-house' between a spoon and a weapon of the chase is a twig or a straw during the capture of ants." (Ibid, p. 76) More generally: "Objects which are interesting, but unpleasant to handle, are forthwith approached by means of a stick" (ibid, p. 79). Compare with this the play at a distance in the child experiments of the preceding paragraph. Further: Should it not be relevant to Nueva's impressive debut in the use of a stick (ibid, p. 32) that her repertoire already contained this: ". . . she scrapes the ground with [the little stick], pushes the banana skins together into a heap . . ."? Even Sultan's extraordinary double-stick solution has its history in play and chance. ". . . [Sultan] plays carelessly with [the two sticks]. While doing this, it happens that he finds himself holding one rod in either hand in such a way that they lie in a straight line; he pushes the thinner one a little way into the opening of the thicker . . ." (ibid, p. 127).[116]

Naturally, manipulations of sticks are to be expected only with animals which can grasp. But the fulfillment of the purely anatomical prerequisites hardly suffices for the occurrence of a problem-solution, if similar performances never happen in the natural life of the species. M. Hertz (11) reports accordingly that ravens, in pulling up a goal hanging on a string, do not resort to their claws, but jackdaws do. For the jackdaw always perches on the booty, i.e., its beak and claws are wont to cooperate, while the raven does not tend to use his claws in this way. Or: clearing away obstacles comes easily to ravens (11, p. 373 f), with difficulty to the chimpanzee (20, p. 65 f.). Why? "With the raven, clearing away . . . plays a rôle in connection with his hiding of things." Ravens, it is known, conceal their treasures by covering them with something. "Uncovering" is thus

[113] From Jan. 11-22, sticks were *not yet* used to reach a goal which had been hung up.

[114] On Jan. 28, in a situation where "Congo" was tied to a pole and the goal was out of reach, a stick which could easily be reached was *not yet* utilized.

[115] On Feb. 2 and 3—in contrast to Jan. 22—sticks *were* used to reach a hanging goal.

[116] "The animals are constantly poking about with straws and small sticks in holes and cracks in their play. . . ." (Ibid., p. 127.)

familiar to them.—Incidentally, the clearing away of obstacles has certain chances of occurrence with all animals, because in the striving for the goal, what is found in the way may easily be inadvertently touched and displaced (1, p. 140 ff).

We can say, in summary, that in order to be found, a solution must offer contact with the instinctive behavior of the species involved.

9. SOLUTIONS INVOLVING EVIDENCE AND EXPERIENCE

In 1-3 above it was pointed out that general experiences about connections everywhere suggest general solution-principles. Now we shall show that even intelligible ground-consequence connections are not entirely independent of experience, inasmuch as it is, for instance, a matter of experience that some such thing as the ground concerned exists at all. Let us take as example the general solution-method: "if it does not work by one way, try another". Without a doubt, it may be everywhere and constantly experienced in our world that different ways lead to the same goal. I find it improbable that, in a world where nothing of this sort could be observed anywhere, it would occur to one to ask himself whether the goal might not be reached by another way. It seems to me just as improbable that, in a world in which alteration and change did not exist, i.e., had not yet been observed, one would think of "altering" what is given if something "other" than the given is desired—and this in spite of all intelligibility of the ground-consequence relation between alteration and "otherness".

Further: It will be familiar to everyone how helpful it is for the solution of mathematical problems to have learned things—again quite apart from all intelligibility. Be it ever so evident, *post festum,* that the existence of a number is proven if one has *constructed* it (see our prime numbers problem, p. 9 f.), still the application of this method of proof is usually the result of learning. Without detriment to their intelligibility, the procedures, for example, of indirect proof, of mathematical induction, etc., can be learned. The same holds for all more special ways of proof as well as for the most general heuristic methods of productive thinking. The importance of learning for such instances of intelligibility is of course connected with the fact that—see p. 61—as a rule the goal may certainly be read off from the solution or solution-method when found, but that, on the other hand, the desired solution-method cannot be read off from the given goal. However, such learning of connection which are totally intelligible *per se* differs very significantly from learning of relatively unintelligible causal relationships, inasmuch as in the former case, from one single experience, the principle can be grasped in all its clarity and precision, so that abstracting induction is entirely unnecessary.—In arithmetic, learning obviously plays an abbreviating part. It is true that each computation is reducible to evident steps,[117] yet an extraordinary amount of time would be wasted, had one not also "learned" the multiplication tables.

10. TOWARDS CLARIFICATION OF THE CONCEPT OF "EXPERIENCE"

The expression that a problem-solution *stems from experience* is so burdened with misunderstandings that it

[117] See p. 58. Thorndike is mistaken when he will see nothing in the multiplication table but an infinity of "connections" or "bonds" (40).

seems well worthwhile to try to make the concept of experience more precise.

1. Not movements and chains of movements are experienced and learned—at all events, in later life—but rather objective ground-consequence structures,[118] whose translation into body-movements presents in general no problem. The notion, stemming from the reflex doctrine and virtually ineradicable, that always and everywhere it is motor impulses which are learned, ought to capitulate completely in view of the fact that every new thing learned (in later life) will from then on occur in thousand-fold variation of the participating movements. (See 20, 29 and 6, for example.)

2. "Experiences" are not in general crystallized in statements, but in attributes of objects. Analogous to the way in which needs become "valences" of things (Lewin), ground-consequence experiences or end-means experiences are embodied in "functional characters" of the objects themselves. The hammer is something for driving nails, the bench is something for resting on (41, also 42, p. 395). More general functional characters of things are, for example, "heavy," "impenetrable," "mobile," "sensitive," "hindering," etc.

3. It has already been emphasized that the experience of "leading to" need not stem from one's own reactions and least of all from blind-reactions. The behavioristic trial-and-error theory has here severely limited the psychologist's field of vision.

4. It has also been pointed out already that not "the same" performance as that now demanded need have been experienced before; rather—because of the possibility of *general* experiences—"similar" performances or causal connections will suffice. Also, this similarity often exists only between the *parts* of the new performance and earlier performances, i.e., the new performance arises through *combination* of several older performances. This possibility has already been indicated on p. 19. In countless solution-processes, a main performance is in this way *combined* with a preparatory performance. After it has once been experienced that something in a definite position or in a definite condition leads to success, this something will sooner or later also be *brought* into that position or into that condition. In animal experiments it can be observed again and again that at first the appropriate position, say, of the box under the goal, of the stick in contact with the goal, is recognized as suitable and thus utilized, and is then later especially produced. The final performance as a whole is thus a combination of two quite different achievements. The performance may also consist in several successive realizations of one and the same solution-procedure. A simple example: If a chimpanzee must first get hold of a long stick by means of a short stick, in order then to reach the banana with the former, this solution-process obviously consists of two "solution-phases" (in the exact sense of the present investigation). The first consists in the demand for a stick sufficiently long to reach the too-distant banana, the second and "supplementary" phase in the demand for a stick sufficiently long to reach this too-distant stick. The two solution-phases therefore represent successive realizations of one and the same solution-procedure.[119]

[118] Compare E. C. Tolman's "means-end-relations" (41).

[119] It would be a mistake, however, to consider a solution easy just because it consists in two applications of one and the same procedure. One finds again and again that the mere inser-

5. It is by no means necessary for the exploitation of experience and of chance that the desired effect or something similar to it have once before been *completely* attained. Frequently it suffices if a mere *approach,* recognisable as such, was somehow brought about. Even the causing of a mere movement of the goal-object may call attention to the possibility of a solution. (1, p. 100 ff, p. 115, p. 117.)

tion of an intermediate goal can make the desired performance extraordinarily more difficult. The mastering of such *wholes with sub-units* requires a considerable "span of the arch of intention" (Beringer). In regard to this span, there are great individual differences. See also K. Gottschaldt (9), chapter VI; further P. Guillaume and J. Meyerson (10): "c'est la *vision d'ensemble* qui manque".

Chapter VI

On Solutions Through Resonance

1. STATEMENT OF THE PROBLEM AND EXAMPLES

In chapter II, 2, it was outlined how a solution can be found, or better, sought out, through the correspondence between the property demanded (anticipated, signalled) and that inherent in what is sought. The process by which, in the perceptual field or in memory (the trace field), an object or a situation is sought out through specific signalling, we then called *resonance*. The "finding of a solution through the resonance effect of a signal" was at that time only mentioned, in order temporarily to be set aside again as a principle which is obvious but much too limited for problem-solving in general. It is the most banal and least rational form of finding solutions, practicable in any world whatever, provided this world contains similarities and repetitions at all. Although the forms of evidence which were discussed in Chapters IV and V permit more rational ways of finding solutions, yet this simple method is by no means made thereby superfluous. Wherever in the course of a solution appropriate means are sought for in the perceptually given problem situation, but above all wherever *previous experiences, learned structures, are included in a solution-process,* resonance effects through appropriate signals may well participate. It is therefore worthwhile to investigate more closely the nature of this finding by resonance.

As a preliminary, let us have a few more examples to supplement those given in Chapter II, 2. I look for a pencil on the table before me. My glance wanders around until it is finally "caught" on the pencil. We must thus distinguish between a signal or "model of search" (such as the approximate representation of the pencil) and a "region of search" (such as the table). Of course the region of search need not be in the perceptual field. Some memory field may be perused instead (with the "inner eye"). It may also happen that a particular region of search is completely lacking. That which is sought (together with its place of abode) then emerges directly "from memory."[120]

"Something of such and such a sort" can be sought just as well as some definite thing. For example, I seek something long like a stick (in order to get hold of an object which has rolled under the cupboard). In this situation I may look around for something suitable in perceptual space, or in some promising region of memory; or something of the sort (a ruler, an umbrella) may occur to me singly, and only afterwards may I ask where such a thing might be and whether it be accessible. One more example: Once I gave to several boys a problem in which they had to find a form of message which could under no circumstances be intercepted by enemies. The first solution-phase or reformulation of the problem—incidentally, analytic of the goal—consisted in that the children search with the model: "something high above," or "something invisible." (For what does

[120] One might be tempted to name a search with definite direction a "vector." However, the direction of such a "vector of search" would by no means be *spatially* defined. This vector would exist, if in a "space" at all, at any rate in some other "space" than, for example, the vector which is defined between the subject and the spatial location of the object when found.

"to intercept" mean?—in any case, to see and to touch.) To these demands corresponded the immanent properties of the proposals: airplane, submarine.

It need not be visual properties which serve as signal of what is sought, unless the search is in a visual field. When I need something with which to drive a nail into the wall, a hammer promptly occurs to me. Thus what is sought here is characterized by its typical function, not by characteristics of its appearance. Only if I look now for the hammer in my environment, does a visual signal come into play.

What is sought need not be a thing. It can also be a procedure, a way to. . . . The signalled property is then: "leading to effect so-and-so."

In such finding of objects and operations, it is of course unnecessary that the demanded or signalled property have at any time already been abstracted from the object. It need inhere in the object solely as a possibility or disposition. Otherwise, the object may occur in my experience for the very first time. It frequently occurs that a familiar object appears one day—under the pressure of a definite seeking—with a "new" property, i.e., one which until then I had never perceived in it. For example, I may look for something to weigh down papers—and a dictionary offers itself. Or the umbrella is suddenly claimed as substitute for a stick. Similarly, an event in my surroundings or in my memory may for the first time be claimed as means-to-the-goal. It is clear that, under the pressure of a definite seeking, very radical *recenterings of the object* can take place (see Wertheimer, 45). Only the "disposition thereto" must be present in the object. Such recenterings of objects may be observed under the very simplest conditions. Let the reader make the following experiment: He is to scan the environment in his room or on the street, putting a stress on red, i.e., he is to seek out everything red which happens to be in the surrounding field. Then in the "relief" of his surroundings, which may be very familiar, an amazing alteration will take place. The structuration is in terms of "red", everything of a red disposition becomes unduly prominent, completely subordinated objects which earlier were hardly or not at all noticed (shop signs, book bindings, neckties) suddenly show off and enter into unsuspected connections with each other. It is not as if one gradually collected red things; rather what is red "leaps to the eye," becomes properly red and dominates the relief. Intellectual recognition limps behind. —For a change, one can also put the stress on "round." Then one suddenly sees quite new forms emerging, and the swarming of red is as though obliterated.

Similar signalling demands on the perceptual field may occasionally occur without our intention, namely, on the basis of so-called *sets* such as, for example, may remain after an intensive occupation with, and a seeking of, objects of a given kind. A personal observation which each time amazes me afresh: if for a while I have intensively read music, it often occurs that for minutes at a time I cannot afterwards look at a so-called homogeneous surface without having all accidental microstructures of this surface at once and inevitably grow into notes. If one realizes how varied are the symbols occurring in musical notation, it will be understood that countless accidental microstructures may be transformed into such symbols. (A larger oblique flaw, for instance, appears spontaneously as two heads of half or whole notes pressed obliquely against each other in the second-interval, or—if it be narrow enough—as a line connecting several notes.) Obviously this is not a case of hallucinations, but of changes in the perceptual field under a completely involuntary pressure in the direction of notation.

The conditions are somewhat different in experiments which K. Gottschaldt reports (8). By presentation of periodic series of visual stimuli, Gottschaldt produced in his Ss at definite positions in the series a specific expectation-pressure, under which a test figure, ambiguous in itself, was perceived in a correspondingly specific way.

2. SUPPLEMENTARY COMMENTS ON THE CONCEPT OF RESONANCE

In order to familiarize ourselves with the concept of resonance as we use it in the present investigation, we may here mention that recognition—individual as well as general[121]—indubitably occurs through resonance. How should a percept acquire even a mere "quality of familiarity," if not by some degree of quasi-resonance of definite memory traces?

By von Kries and Becher it has already been conclusively demonstrated that the fact of recognition is not compatible with the "conduction hypothesis," according to which a complex of sensations and its trace must be coordinated by specific pathways. If a particular object is recognized, although its retinal image lies now here, now there—moreover, in different sizes and perspectives—then a specific conducting track between retinal location and locus of trace cannot possibly be responsible for the excitation of the appropriate trace. For everything must be spatially (neurally) connected with everything else, since from every part of the retina every trace can be aroused. The selective principle therefore cannot consist in specific neural connection. It must be supplied through the relation of the properties of the externally caused excitation-pattern and of the trace. For the time being, no more than this is meant by the expression "resonance."

The process of recognition can be crudely represented as follows: a peripheral stimulation sends into the nervous system a particular wave of excitation (specific to its content, perhaps very complicated), and to this wave kindred traces respond as though by resonance. Under these circumstances the neural pathways would practically play the rôle of an indifferent, impartial medium.

According to W. Köhler's theory,[122] the reproductive effect of similarity is to be conceived as pair-formation on the ground of the gestalt-factor of similarity in its application to the time-dimension. The experiments carried out by H. von Restorff give interesting information about the dependence of recall on the temporally intermediate field. In very brief summary: *Ceteris paribus,* a reproduction or a recognition takes place more easily across an intervening period filled with material of different nature than across one filled with similar material. Therefore, our further statements on recall by similarity are to be understood with the additional condition: under appropriate circumstances as to the intervening field.

3. DISCUSSION OF SOME EXPERIMENTS

The subject is to say what this is: "red, round—juicy—soft." After red and round, one subject reproduced "rubber

[121] Individual recognition: I recognize a house in which I lived years ago, a favorite song hit, my coat. General recognition: I recognize houses as houses, songs as songs, coats as coats. Recognition in this latter sense occurs constantly.

[122] [See (23), Chapter III, p. 125 f.]

ball," after juicy, "apple," after "soft" first "plum," then "tomato."—A second experiment: "long, pointed, cold." (An icile is meant.) The Ss name at first things like "knitting needle," "sword." The experimenter then adds to the signal the property "brittle." Now the answer is more likely to be "icicle." If the model of search is further enriched by the property "hanging," the solution "icicle" seldom fails to appear. What have we here?

1. The individual properties given do not act "as pieces," as an aggregate,[123] as if each aroused its own trace, and as if then by superposition the trace common to a maximum of individual properties, or the trace preferred for another reason, were selected. Rather it can be directly shown as a phenomenal fact that the individual properties are joined into one unified "model of the thing," each in its specific position in the structure. In other words, the model of search "constructed" out of the given properties arouses as a whole appropriate trace-wholes. The process resembles, in this respect, the recognition of a perceptual thing, where again not individual properties arouse individually adequate traces, but the concretely structured thing as a whole establishes contact with a corresponding trace-whole.

2. But the model of search is no complete, sharply defined "thing." It is a partially undetermined, vague, fluctuating (or purposely varied) thing-schema with "gaps" which may be filled provisionally.

3. The model of search presupposes an understanding of the instruction, and this involves two phases: (1) Every single one of the verbally given properties must first be understood *per se*. This understanding occurs through arousal of the traces corresponding to the particular genus of property, by way of the proper word-traces. (2) From the aroused single traces, the model of search with its character of a unified entity is then constructed through a process of "structural combination" (see 4 below).

4. STRUCTURAL COMBINATION

The just-mentioned principle of structural combination has fundamental significance for the psychology of the understanding of language.[124] If someone says, "Imagine a yellow crow on the table there, with a cigar in its beak," a unitary image arises—after more or less effort. The elements and the relations (the abstract structural loci) are all known. But their combination is *new*. A crow and a table can be brought into the relation "on;" a cigar and a crow's beak can be brought into the relation "in." The structural locus, "color of the crow," can be filled with yellow, etc.—New combinations are obviously not simply accumulations, but rather completions of structures. Elements are set in relations. Each "application" of general statements, general experiences to new material occurs through structural combination. Of course, it is not the traces themselves which participate in such a combination, but rather actual processes which are "derived" from the former.

The general law of structural combinations runs as follows: Excitations of any two traces, ideas, or images, or of an idea and a percept, may actually be put into any relation whatever into which (according to their nature) they can enter as related terms. Combinations can

[123] See the "constellation theory" of thinking proposed by W. James, G. E. Müller, and Poppelreuter among others. This theory was refuted first by Selz theoretically (36, p. 281 f.) and by Shepard and Fogelsonger experimentally (38), but still leads a somewhat vague existence (especially among psychiatrists). Cf. the alleged steering of the reproductive process by so-called "dominant ideas" (*Obervorstellungen*).

[124] See in this connection (36), p. 162 ff.

be conditioned from within (compare the free play of imagination) as well as determined from without (verbally). Whoever reads a noval, carries out from the first to the last page verbally determined combinations. Incidentally, combination is nothing but a special case of "restructuration of material" in general.

5. TRANSFORMATION INTO CONCRETE MODELS OF SEARCH

Comprehended instructions are frequently not yet proper models of search. The original characterization of what is sought, the original signal, must first be suitably transformed. A few observations concerning such tranformations follow. They come from problems of this type: "Give the name of a town in which the fourth letter from the end is *t*." With such problems, the original signal, which is recognizable in a concrete name only by special examination, and which therefore "hits" no directly perceptual character of this name, has to be changed into a concrete perceptual model of search. This fact is characteristic of resonance.

Some Ss, given the problem above, proceeded as follows: they formed a model of search of the type: T-/--,[125] with which frequently a definite first vowel was anticipated. With one S this led to the solutions (all not quite right): Tegel, Togo; then, after the two-syllabic schema was given up as "too long:" T(h)orn. Or sometimes a name of the form: __/stein or __/tern was anticipated, at which appeared: Gastein, Kaiserslautern.

One sees from the examples given how a quasi-perceptual (and on this account rather specific) model arises, which organises itself rhythmically around some outstanding elements. Such a model is incomplete in the sense that it contains "gaps".

6. "PRÄGNANZ" OF A SIGNAL

We saw that not every "corresponding" of signal and object already makes possible a resonance of the object to the signal. The signal must "suit" the object "well," it must in this sense be "prägnant." There are several conditions for this:

1. Completeness, wealth of content of the signal. A simple example for this would be the fact that the trace "icicle" is the more easily aroused, the more of these properties: long, pointed, cold, brittle, hanging . . . are anticipated, or signalled.

2. Succinctness and conciseness of the signal. At best, no properties should with any weight participate in a signal, which the object to be found may, but need not, exhibit. An example: I seek something with which this brass nail may be driven into the wall, to support this Van Gogh portrait. Such a signal certainly corresponds to "hammer", but obviously it does not matter to the hammer whether it is meant, just here and now, to drive just a brass nail into the wall, for just this purpose. The finding of this solution therefore necessarily takes place in abstraction from certain superfluous circumstances. Without such an abstraction, pertinent general experiences could never be made use of at all.

3. Correctness of the signal. The signal should not be "false," i.e., it should not contain anything which the object not only does not contain, but which it contains in another version. Example: I look around for a blue mail box —in vain, for in this country the mail boxes are dark green.

A search often finds itself in the following characteristic conflict situa-

[125] The short level signs signify letters, the slanting one signifies syllabic division.

tion. Let objects be sought which have the property *b*. Let there exist one class of objects with the properties *bc*, another class with *bd*, a third with *be*, etc. Now, if I enrich my model of search by one of the additional qualities, e.g., *c*, the chances for *bc*-objects certainly increase

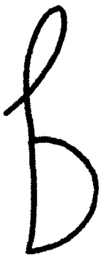


FIG. 22a

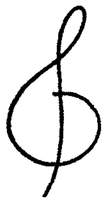


FIG. 22b

(on account of the greater completeness of the signal). But at the same time the chances decrease for *bd*- and *be*-objects, with regard to which the anticipation is now incorrect. Whether I therefore attain more with the model *bc* than with the model *b* will depend, among other things, on how large is the class *bc* in relation to the rest.

4. "Perceptual adequacy" of the signal. The signal: "names of towns of which the fourth letter from the last is *t*"

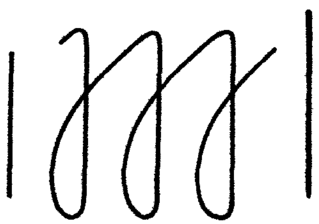


FIG. 22C FIG. 22d

was not in this sense adequate. It was correct, but it did not point to a perceptual characteristic. As further examples some facts can serve here which W. Köhler urged against G. E. Müller's thesis of the decisive significance of strong cohesion in a complex not only for association but also for reproducibility (19). From the "part" in Fig. 22a-e there issues no marked tendency to recall the rest of the violin clef in Fig. 22b; because that "part" is no characteristic, no genuine part of the whole. For the same reason, the whole complex in Fig. 22c will more readily be reproduced from Fig. 22d than from Fig 22e.[126]

Of the four conditions enumerated, the "general law of anticipation" of Ach and Selz involves only the first, that of completeness.

7. ON REGIONS OF SEARCH

As a rule, a search is directed to something which, by its nature, is or can be part of a "field" or of a "concrete region" of some kind. This makes it possible for

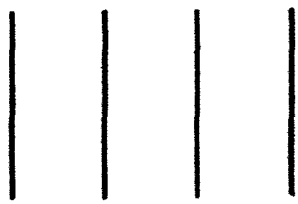


FIG. 22e

the model of search to be applied directly to the specific field, the specific region—in limiting cases, only to a minimal region which happens to be touched in hasty perusal. (See the examples on p. 75f.)

Thus, in the experiment with names of towns, German names were sought, whereat frequently more particular regions within Germany were "perused." A subject who searched with the schema

126 These figures are taken from W. Köhler's "Gestalt Psychology" (21).

"___/stein" was clearly directed to the mountainous southwest of Germany. He had the feeling that more names ending in ——stein must exist there. With another problem: "four-syllabled towns," one S began at once to search in the United States, whereat he repeatedly perused the map as he visualized it.

Similar observations may be made during that familiar parlor game: to name, for example, great men whose names begin with M. Here again, not only more concrete models of search are constructed (e.g., Mi——), but also regions are demarcated and perused (e.g., "physicists" or "French writers" or the like). These regions are understood as concretely as possible, i.e., not as abstract classifications but, say, as a certain vicinity on the bookshelf, or as a definite historical volume, or the like. Such concrete regions are then more or less systematically perused whenever necessary, i.e., in case what is suitable does not at once "spring to meet" the model of search.

In other words, the properties of the original signal which refer to content are supplemented by *"topical"* aspects.

The following modification of our icicle problem shows how the determination of a model of search in terms of content may be supplemented or even replaced by a determination in terms of region. If instead of the properties "brittle" and "hanging" one gives to the S the regional determination "winter", one does for him (or for the solution, "icicle") about the same service.

The cooperation of determination by content and spontaneous determination by region is illustrated in the following experimental protocol: Problem: "round, brown, heavy"?

Reactions: (1) "Chestnut (but not heavy enough)." The feeling now arises that "some such thing could exist elsewhere in the plant kingdom." A garden scene is perused in thought. Thence arises (2) "gourd." Then the plant kingdom is given up. After a while there appears (3) "head with brown hair," then, with direction toward the animal kingdom, (4) "a brown animal rolled up, a bear." Now once more a change of region results: an artifact is sought. A furniture warehouse appears promising and is perused. In connection with this (after the spatial independence of the object has intentionally been given up), finally (5) "a spherical foot of a wardrobe" or a "decorative globe" on a newel post is found. Then this region as well is given up again. The initiative is gradually exhausted. (6) "Puff-ball" and (7) "boulder" are still named, but in both cases the requirement "brown" is felt not to be well enough fulfilled.

A region of search can be found in various ways. Often, as the characteristic "sphere" of a certain object, it is simply aroused together with the latter (e.g., a garden as the sphere of a botanical object) and then serves as region for further search. Or the arousal is mediated by a signal expressly directed to a pertinent region of search (see 9 below on "co-arousal").

The search within regions is generally carried out through *perusals*. Thus spatial fields (a landscape, a book) or temporal fields (a day or a whole section of life) may be perused in a more or less cursory fashion. Perusals of regions appear especially when no "prägnant" model of search has been found, i.e., none which has been able directly to "attract" a suitable object out of the larger field. It has already been mentioned that determination in terms of content may be partially replaced by regional determinations. In the limiting case, where no model of search in terms of contents is present at all, but only a region of search and some signal in terms of pure function: "something to . . .," nothing but perusal is left. The object is then not found through its suitability at all, rather the search stumbles on it en route, and establishes its suitability, its fitness, only subsequently by special examination. But this means merely that the search has become purely "topical."

8. TOPICAL ANTICIPATION

How does the answer arise to questions of the type: "What is *on the left of* that inkwell?" "What country lies *between* Germany and Spain?" "*With what* does one drive a nail into the wall?" "What is *the color of* ripe tomatoes?"

In distinction to the earlier problems of this chapter, not "properties" (in the narrower sense of the word) are now anticipated, but "relations," "gaps," more precisely: *"places in relation structures."* Therefore we call this *"topical anticipation"* (xRb, i.e., x in the relation R to b).[127]

The finding of an object which is thus topically anticipated we shall call "reading off," because what is sought is "read off" from a definite structural locus. The whole process of finding is carried out in two steps. First, the explicitly signalled objects are aroused, to which the object sought should stand in the relations R_1, R_2. . . . And those objects then become parts of the signal by which what is sought will itself be found. *This decisive signal therefore already contains real parts of just the individual perceptual or trace complexes to which also belongs what is sought.* This is the essential difference from an anticipation purely in terms of content. (We saw however—cf. 7—that an anticipation in terms of content generally tends to involve topical anticipation in that it operates through perusal of particular regions.)

What we called insight in reading off (Chapter IV) is throughout of the type of such a topical finding, although complicated by certain additional conditions, which were discussed there. In such a reading off as: "What relation exists between a and c, if $a>b$ and $b>c$?," the finding is in fact mediated by the locus: "between a and c". The relation sought is read off from this locus.

There are topical anticipations in which the relation R indicates a perusable direction, i.e., a direction in which—from the starting object b—the glance may shift until it hits on something situated in this direction. Such is the case, e.g., with the question: "What stands on the left of that inkwell?" Here the glance can wander leftwards from the inkwell, till it is, so to speak, caught on something "standing" there. Such is, however, no longer the case with problems like "solution of" or "the smallest of;" for here no simple direction of perusal corresponds to the relation.

The case of topical anticipation most important for us is that in which "something that leads to . . . ." or "something to . . ." is sought. Since, as we saw, a concrete aspect of experience corresponds to the "leading to" (see Chapter V, 2, 4 and 5), from each experience of a "leading to", that which leads-to can in fact be "read off" by resonance, perhaps in conjunction with a certain re-perusal of the temporal course of events.

Frequently the structural locus of an object has long since become a "quasi-property." The object has so to speak absorbed its typical locus, so that the locus now inheres in it as a "property," as content. Thus, for instance, the functional locus of the hammer in the complex of driving a nail has long since

[127] In a further, but far too formal sense of "topical", it means of course also topical anticipation if the object is signalled by means of definite "properties". For in a way a "carrier" of such properties is what is looked for (thus xRb, where b is the properties, R the relation of carrying).

become a "property" of the hammer (p. 73).

9. RECALL AND "CO-AROUSAL"

In the usual reproduction by similarity, not only "what is similar" is aroused, but "what is connected with it" is *"co-aroused."* In general, only what is co-aroused enters into consciousness—which accounts for the fact that for so long in the theory of association "one did not see the similarity-factor for so much contiguity."[128] If we call a' the reproduction stimulus—which need not be a model of search—and if we call a the whole complex which is reproduced, then there is always a "difference" a-a', of which one can say that it is "co-aroused." If a-a' consists only of dependent properties (for instance, the glassy transparency of the icicle which is anticipated as something long, cold, pointed), one will speak of co-arousal merely for the sake of theoretical consistency. For the most part one speaks in such cases simply of "reproduction by similarity", thereby ignoring the accompanying co-arousal. To the degree, however, to which a-a' is given as a relatively independent and emphasized part of a (see the frozen gutter on the roof, the house, the whole winter landscape in the icicle example), the expression "co-arousal" appears more natural. With this sense of the term "co-arousal", in recall often "the jacket is sewn to the button". Whole atmospheres, wide regions in space and time can be "co-aroused". Under these circumstances one usually speaks of "recall by contiguity".

[128] Consequently, it was believed that all recall by similarity must be reduced to reproduction by contiguity. No one before Höffding saw the circle: if a p' is to reproduce a q, it must first of all reproduce that p with which the q is in its turn associated. The reproduction by contiguity $p' \rightarrow q$ is therefore necessarily preceded by a reproduction by similarity (a resonance event) $p' \rightarrow p$.

Now how does co-arousal come about? Two forms of co-arousal are to be distinguished:

1) A difference a-a' can be co-aroused simply because it belongs (in a dynamic, not a logical sense) inseparably to the totality of a. In other words, a-a' is aroused because a can be aroused only as a whole, because the common factor of a and a' is dynamically dependent within a. Thus, whether and how strongly something is co-aroused depends on the intimacy and firmness of the figural coherence within a. This form of co-arousal may be based on the gestalt factors of "proximity" and "closure". "While selection on the ground of the gestalt factor of similarity occurs between a "new" process and a trace, the further reproductive selection on the ground of proximity and closure takes place not between a "new" process and a trace, but between an aroused trace and another which is "near" or which "closes" the whole in question.

Incidentally, the classical laws of association refer throughout to certain very special conditions which, among others, determine the formation, the kind of coherence and the degree of stability of trace-systems.

We shall designate this form of co-arousal as "co-arousal of dynamically dependent parts within the system a that contains the counterpart of a' ".

2) But a co-arousal may also be anticipated, which means that what will be co-aroused is topically anticipated. If something occurs to me on the ground of a similarity, I can intentionally "spin it out", supplement it in certain directions beyond its own locus. (Compare Selz' concept of "determined complex-supplementation".) But I can also restructure its interior. (Compare Selz' concept of the "determined abstraction".) The mere question: "What occurs to you

in this connection?", is already a topical anticipation, even though a very vague one.

With great emphasis Selz has pointed to the existence of such "determined" reproductions (in contrast to "diffuse" ones). He discovers them even among the most meaningless performances of memory. (Cf: "What is to be recalled can in such cases . . . be anticipated in abstract form: as the first syllable of the learned series, or the syllable following the previous one . . ." (36, p. 289).) However, Selz seems to me to go decidedly too far, when he regards all involuntary reproductions (co-arousals) as determined, i.e., as anticipated. Undoubtedly, there is such a thing as our simple "co-arousal of dynamically dependent parts within the system *a* that contains the counterpart of *a'* ". To speak even here of anticipations is to dilute the concept. Passive tolerance is no longer anticipation.

10. PARTIAL RECALL

It would be erroneous to assume that a certain trace must be aroused completely or not at all. Rather, a trace contains as a rule a multiplicity of features and references, of which a few may very well be recalled separately from the others and "before" them. Now it very often occurs in the search for something *de facto* familiar that the corresponding trace "projects" in advance characters which were not directly anticipated, that, one might say, it "foreshadows itself". Both parts of the content and topical factors may be "foreshadowed," i.e., other components of the whole trace to which that *x* belongs, in their characteristic relation to *x*. Such "projections" from what is sought are now incorporated into the existing anticipation and "transform" it.[129]

Here is an example. Problem: name of the author of *A Hero of Our Time?* Actually, I seek "that three-syllabled name ending in -ow". Final solution: Lermontow. In many processes, only some such thing as "Well, I know that!" precedes the final solution. This is probably the weakest "projection" which a trace can make at all.[130]

Through this "pre-arousal" of other parts of the trace-whole to which a subject is referred, important changes in the meaning of what is sought may be caused. For instance, someone searches his memory for the date of a journey. Then it occurs to him: "But that was on the day after the . . . concert in the Philharmonic Hall, and that was . . . ?" The missing date of the journey thus undergoes a *recentering* into the date of the day after the concert. Or: One seeks the author of a certain book, and finds himself looking for—the man who has also written such-and-such other books. In ordinary attempts to remember such recenterings of the model of search are carried out everywhere.

[129] In such cases, Selz speaks of "successive actualization of knowledge" (36), p. 45 and p. 62 ff.

[130] A further excellent example from Lichtenberg's notes: "When, on the 24th and 25th of January, 1790, I tried to recall the name of the Swedish writer and bookseller Gjörwell, which apparently I could not find at all, I noted the following: At first I despaired completely of ever finding it by myself. After a little while I observed that when I pronounced certain Swedish names, any approach to the right one was dimly felt as such by me. Indeed, I thought I noted when I was nearest to it. And yet I suddenly lost it and seemed again to feel that I should not find it at all. What a strange relation of a lost word to the others which I still had with me, and to my head! By the way, I always gave preference to two-syllabled names. Also Bjelke, Njökoping and the like were the nearest, on account of the *ö* and the *j*. Finally, after I had tortured myself the night through, and had certainly aggravated my nervous attacks thereby, I endeavored to find the initial letter; and after I reached G in the alphabet, I stopped short and at once said Gjörwell; . . ." (From *Aphorismen und Schriften*, published by E. Vincent, Verlag Kröner, p. 25.)

PART THREE

FIXEDNESS OF THOUGHT-MATERIAL

CHAPTER VII

ON FUNCTIONAL FIXEDNESS OF REAL SOLUTION-OBJECTS[131]

1. SETTING OF THE PROBLEM. THE CONCEPT OF HETEROGENEOUS FUNCTIONAL FIXEDNESS

IN CHAPTER II, 6 and 7, it was pointed out that the different parts of the situation, whose (appropriate) variations represent solutions of the problem, or *which enter into* solutions as "material", may display very different degrees of "disposability" (looseness). For the psychology of thinking, there hardly exist more fundamental differences among the various relevant elements of a problem-situation than those which determine how easily or with what difficulty they may be recognized as conflict-elements or as solution material. These differences are independent of possible "knowledge" by which *post festum*—the elements concerned could be evaluated with respect to their conflict-character or their suitability as material.

A few of the factors which determine disposability, specifically that of conflict elements, have already been worked out in Chapter II. Now we shall examine more closely the *disposability of solution material,* in the more specific form of "real solution-objects sought".

Whether a sought "object" is found more easily or with more difficulty depends, among other things, on the degree of *"fixedness"* of the object. A chimpanzee who stands in need of a stick (something long, firm . . .) sometimes has difficulties in recognizing the stick in a branch still growing on the tree, in seeing it as a percept apart (20, p. 106). On the tree it is a "branch", a part of the visual figural unit "tree", and this part-character—more generally, this *"fixedness"*—is clearly responsible for the fact that to a search for something like a stick, the branch on the tree is less "within reach" than the branch on the ground.

What we just named "fixedness" may, however, be conditioned *functionally* as well as by such factors of visual organization. For instance, a stick that has just been used as a ruler is less likely to appear as a tool for other purposes than it would normally be. In the following, the discussion will be chiefly of such functional fixedness ("bias"), more particularly, of *heterogeneous* functional fixedness, i.e., fixedness as the result of a function *dissimilar* to that demanded. The question is: *What determines whether, and to what degree, heterogeneous functional fixedness of an object hinders the finding of this object?*

On this question I undertook a series of experiments.[132] The principle was as follows: For a particular purpose, a certain function, a suitable object is needed. *This object has already been used in the same problem-context, but in another way, in another function.* Question: what effect has this pre-utilization? When does it hinder the selection of the object

[131] The minuteness of detail in the following treatment of a special problem is somewhat out of proportion to its theoretical importance in the framework of the present investigation.

[132] For the conscientious carrying out of these experiments, I am greatly indebted to Miss Rosenbusch, cand. phil.

for the new function, the *"recentering"* of the object?

Be it expressly noted that what, in the present chapter, is stated for thing-objects (specifically tools) is valid, in principle, for thought-material in general. (See for instance, Chapter VIII.)

2. EXPERIMENTAL PROCEDURE, METHOD OF EVALUATION, AND PROBLEMS

We experimented with all sorts of objects in daily use (e.g., boxes, pliers, etc.), which were first claimed in their usual function (F_1) and then, within the same problem-situation, for a new, unusual function (F_2). The crucial object was each time to be selected as the suitable tool out of a great number of objects which lay in confusion on a table.

In our problems, the pre-utilization of the crucial object was chosen in such a way as not to give it a special prominence in the problem situation. In other words, in F_1 no new centering took place, but solely a freshening, an "actualization" of the usual centering of the object concerned. For F_2, on the contrary, the object concerned was "unprepared," although by no means inappropriate.

In order to observe the effect of fixedness on recentering, each problem was given in two settings, once without and once after pre-utilization of the crucial object. The setting without pre-utilization we shall briefly designate *w.p.*, that after pre-utilization, *a.p.* The most important experiments were carried out on five different problems. One-half of the Ss received the problems in the settings: 1) w.p.; 2) a.p.; 3) w.p.; 4) a.p.; 5) w.p.; the other half of the Ss, in the opposite settings. In this way, differences of results in the w.p. and the a.p. experiments were made independent of individual differences among the Ss and among the problems.

The following is a short description of the five problems and of the experimental technique.

The "gimlet problem:" Three cords are to be hung side by side from a wooden ledge ("for experiments on space perception"). On the table lie, among many other objects, two short screw-hooks and the crucial object: a gimlet. *Solution:* for hanging the third cord, the gimlet is used. In the setting a.p., the holes for the screws had yet to be bored; in w.p., the holes were already there. Thus, F_1: "gimlet"; F_2: "thing from which to hang a cord".

The "box problem:" On the door, at the height of the eyes, three small candles are to be put side by side ("for visual experiments"). On the table lie, among many other objects, a few tacks and the crucial objects: three little pasteboard boxes (about the size of an ordinary matchbox, differing somewhat in form and color and put in different places). *Solution:* with a tack apiece, the three boxes are fastened to the door, each to serve as platform for a candle. In the setting a.p., the three boxes were filled with experimental material: in one there were several thin little candles, tacks in another, and matches in the third. In w.p., the three boxes were empty. Thus F_1: "container"; F_2: "platform" (on which to set things).

The "pliers problem:" A board (perhaps 8 inches broad) is to be made firm on two supports (as "flower stand or the like"). On the table lie, among other things, two iron joints (for fastening bars and the like on stands), a wooden bar perhaps 8 inches long (as the one "support") and the crucial object: the pliers. *Solution:* this pair of pliers is utilized as the second support of the board. In the setting a.p., the bar was nailed to the board and had to be freed with the help of the pliers; in w.p., it was only tied to the board. Thus F_1: "pliers"; F_2: "support".

The "weight problem:" A pendulum, consisting of a cord and a weight, is to be hung from a nail ("for experiments on motion"). To this end, the nail must be driven into the wall. On the table lies, among other things, the crucial object: a weight. *Solution:* with this weight (as "hammer"), the nail is driven into the wall. In the setting a.p., the weight is given expressly as pendulum-weight (with the string already tied to it); in w.p., a joint serves as pendulum-weight. Thus F_1: "pendulum-weight"; F_2: "hammer".

The "paperclip problem:" A piece of white cardboard with four black squares fastened to it is to be hung on an eyelet screwed into the low ceiling ("for visual experiments"). On the table lie paperclips, among other things. *Solution:* a paperclip is unbent, one end is fastened to the eyelet and the other put through the cardboard. In the setting a.p., the four black squares must previously be attached to the cardboard with paperclips; in w.p., on the other hand, they must be glued to it. Thus F_1: "something for affixing"; F_2: (unbent) "hook".

The differences among the five problems are to be discussed later; see 6.

The general *instruction* for all the problems ran as follows: "You will receive several little technical tasks. For solution, certain objects are needed which you will find among the objects here on the table. Everything which lies on the table is completely at your disposal. You may use what you like in any fashion you wish. Please think aloud during the experiment, so that I may hear as many of your ideas as possible, including those which you take less seriously".

With each problem there lay on the table—aside from the objects already mentioned—all kinds of material, partly less suitable and partly completely unsuitable for the solution, such as paperclips, pieces of paper, string, pencils, tinfoil, old parts of apparatus, ashtrays, joints, pieces of wood, etc. Each problem had its own inventory. (No object was put at the subject's disposal which might be better suited to the solution than the object then crucial.) The objects lay in apparent confusion, but in definite places. The crucial object never occupied a prominent place.

The experiments were *evaluated* in two ways: (1) The solved and the unsolved problems were counted. Of course, a problem counted as "correctly" solved only when it was solved by use of the crucial object, which, as stated, was always the best and simplest of the possible solutions. A problem was broken off as unsolved if for two to three minutes the S produced no more proposals, and if at the same time his attitude had become so negative that no more sensible ideas seemed forthcoming. (2) The proposals preceding the solution and different from it, the "pre-solutions", were counted (but only with those experiments in which the correct solution was finally found, as otherwise measurements 1 and 2 would not have been independent of each other). As "presolutions" counted not only those actually carried out, but also proposals merely formulated, also such as the S rejected as unsuitable. If, however, an object was only "grazed", i.e., just touched or picked up quite briefly and silently laid aside again, the fact did not count as a pre-solution.

Of the two methods of evaluation just described, the first is naturally the more adequate and by far the more important, while the second is rather superficial and dependent on chance influences. We shall find, however, that both methods yield results which are essentially in agreement.

3. PRINCIPAL EXPERIMENTS AND PRINCIPAL RESULTS

The principal result of the experiments is immediately evident from Table 3.

We see that the results of the a.p. experiments clearly deviate from those of the w.p. experiments in the expected direction. This holds in both measurements, which are independent of each other, and not only for the average of all five problems, but also within each single problem. Only in the weight problem are the two averages of pre-solutions equal.

Therefore we can say: *Under our experimental conditions, the object which is not fixed is almost twice as easily found as the object which is fixed.*

The quantitative results were supported and clarified through qualitative findings. When, at the close of an a.p. experiment, the S was asked: "Why have you not used this object" (the crucial one) or, "Why have you used it only so late?", the answer was frequently: "But that is a tool", or, "Such a use would not be suited to the material", or, "I thought it was there simply for . . . (F_1)".

TABLE 3

Problems		No. of Ss	No. of problems solved	No. of problems solved in %	Average no. of pre-solutions per problem
w.p.	Gimlet	10	10	100	0.3
	Box	7	7	100	1.3
	Pliers	15	15	100	1.9
	Weight	12	12	100	0.8
	Paperclip	7	6	85.7	0.8
Arith. Mean		—	—	97.1	1.0
a.p.	Gimlet	14*	10	71.4	1.6
	Box	7	3	42.9	2.3
	Pliers	9*	4	44.4	2.3
	Weight	12	9	75.0	0.8
	Paperclip	7	4	57.1	1.5
Arith. Mean		—	—	58.2	1.7

* The inequalities in the number of Ss in w.p. and a.p. are due to the fact that certain Ss transformed the problem-setting intended for them into the opposite setting. In the gimlet problem, for instance, three Ss assigned to the w.p. group actually had to be counted in the a.p. group: one attempted, using the gimlet, to stuff the cords into the holes which were already there; the other two bored holes with it because they did not quite trust the holes which were there. On the other hand, one S assigned to the a.p. group immediately picked up the gimlet as "thing with which to hang up. . . ". Thus he did not use it in F_1, and had therefore to be counted with the w.p. group.—In the pliers problem, three Ss did not utilize the pliers for freeing the bar which was nailed to the board, and therefore had to be counted in the w.p. group.

The last observation might suggest the following objection: It is not the effect of a "bias" of the crucial object which is measured in the experiments but rather the effect of a bias of the subject. The S may be of the opinion that the experimenter has put the crucial object on the table especially as a tool for F_1, that it does not belong to the actual experimental material. (Such false "self-instructions" are not infrequent in the relatively artificial situations of the laboratory.) This objection, however, hardly holds water. In the first place, little significance should be ascribed to statements after the fact, such as "I thought. . . ." They often express only "rationalizations". Secondly, there were many among our Ss for whom it was as if "the scales had fallen from their eyes" when the crucial object was afterwards pointed out. They did not have the feeling of having been victims of a false interpretation of the experimental conditions. In the third place, certain experiments to be cited later (see p. 96 below) also refute this objection.

4. ON FIXEDNESS "RELEVANT TO THE SITUATION" AND ON "CONTACT"

With the box problem and one not mentioned till now, the cork problem, we undertook a few more specific experiments on the possible influence of certain differences in the way the crucial object was pre-utilized.

The cork problem consisted of the following: A triangle was to be drawn on a piece of cardboard which was in turn to be fastened to a wooden bar. The wooden bar was then to be fixed in a doorframe without the help of nails. But the bar was about 2 cm. shorter than the distance between the two sides of the frame. On the table lay, among other things, the crucial object: a cork. Solution: With the help of the cork, the bar is wedged between the sides of the frame. In the a.p. setting, the cork stuck as a stopper in an ink bottle, from which the ink for drawing the triangle was to be taken. In w.p., the cork lay free on the table, at some distance from the ink bottle. Thus, F_1: "stopper;" F_2: "thing for wedging something."

The cork problem is clearly related to the box problem even by external appearance. Here, as there, F_1 is given not by an action but "statically." Here, as there, the crucial object is fixed by F_1 not only functionally, but visually as well. The cork problem, incidentally, proved to be the most difficult of our six problems.

These two problems were also given in the following variation (a.p.′): The crucial object was burdened with a *function F_1 "peripheral to the situation,"* not with one *"relevant"* or *"central to the situation."* That is, in the box problem, the three boxes were filled not with material relevant to the problem (candles, matches, tacks), but with neutral material (buttons and the like). In the cork problem, analogously, the cork was not stuck in a full ink bottle necessary to the drawing, but in a superfluous empty one (the triangle was here to be drawn with pencil). The crucial objects, therefore, had a function F_1, it is true (therefore the designation a.p.′), but a relatively peripheral, irrelevant function F_1. We expected that in this situation the recentering would succeed more easily, in the belief that an irrelevant ("dead") function comes nearer to absence of function than does a "living" one. The experimental results (see Table 4), and above all certain qualitative remarks of the Ss, soon made us aware, however, that in this reasoning we had obviously overlooked an important factor.

TABLE 4

Problems		No. of Ss	No. of problems solved	No. of problems solved in %	Average no. of pre-solutions per problem solved
a.p.	Box	7	3	42.9	2.3
	Cork	7	1	14.3	2
a.p.′	Box	7	1	14.3	4
	Cork	7	0	0	—

11 of the 13 Ss who had not solved the cork problem in the a.p. and the a.p.′ settings subsequently received this same problem in the w.p. setting. Under these circumstances 8 of the 11 Ss, i.e., 72.7% solved the problem. Of course, the fact has here been of some influence that the Ss were especially directed to the stopper by the difference: "formerly stopper in full ink bottle, now stopper near empty ink bottle" (although we tried to counteract this suggestive difference by other differences introduced *ad hoc*). The suggestive effect of such differences is considerable.

Table 4 shows—and similar experiments undertaken in a seminar furnished corroboration—that, far from facilitating the solution, the a.p.′ setting actually makes it more difficult in comparison with the a.p. setting. For this reversal of the expected results, a factor is responsible which was also met elsewhere: *the more central to the situation is F_1, the greater the "contact" between the S and the crucial object.* For example, the candles, as the most important objects in the box problem, are undeniably in the center of the material offered, one might almost say: in the fixation point. And the box holding the candles profits from this. For, it is mostly emptied first, i.e., before the two other, less central boxes. The box of buttons, on the other hand, is a quite peripheral part of the problem. In the cork problem, the situation is to some degree similar. The same factor of "contact" was expressed in the remarks of two Ss when they solved the a.p. setting of the paperclip problem: "in handling the paperclips" (in affixing the square, thus in F_1), "I became aware of them."

"Contact" between S and object probably played a rôle also in the following proposal from the pliers experiment: "Break off a piece of the board and use it as the second support." The board stands, of course, in the center of the situation—the S even had it in his hand—just as did the candles in the box experiment. (Of course, the fact also has an influence here that the board can "give up parts;" further that, like the first support, it is wooden.)

Everything indicates that this factor of "contact" has sometimes an effect of facilitating the solution. "Contact" might well be able to overcome the necessarily accompanying fact that a F, which is relevant to the situation causes strong fixation.

5. CORRELATION OF QUANTITATIVE AND QUALITATIVE FINDINGS

Until now, the experimental results have been considered only in reference to the general difference between the w.p. and the a.p. settings. However, it is obvious that the different problems exhibit the hindering effect of heterogeneous fixedness of the crucial object in different degrees. The box problem shows the greatest difference between the w.p. and the a.p. settings, the weight problem the least.

If one orders the five problems according to the size of this difference, measured by its most important indicator, namely, the size of the quotient:

$$\frac{\text{no. of solutions with w.p.}}{\text{no. of solutions with a.p.}},$$

abbreviated: $\frac{S(\text{w.p.})}{S(\text{a.p.})}$, the rank order in Table 5 results:

TABLE 5

	Box	Pliers	Paperclip	Gimlet	Weight
$\frac{S(\text{w.p.})}{S(\text{a.p.})}$	2.3	2.3	1.5	1.4	1.3

In the following discussion, we shall include the cork problem as well. This problem was surprisingly seldom solved in the a.p. setting; see Table 4. The w.p. setting of this problem was given only subsequently, it is true, i.e., after the a.p. or the a.p.′ setting. Therefore the uncommonly large quotient $\frac{S(\text{w.p.})}{S(\text{a.p.})} = \frac{72.7}{14.3} =$ 5.1 is not fully analogous to the quotients of Table 5, and probably too large in comparison with them. Yet an entirely analogous quotient would probably still be among the largest.

Let us now examine these characteristic differences of the quotients in their relation to the psychological nature of the problems concerned. To this end we shall investigate each one of the six a.p. problems as to all the factors which might be supposed to hinder the required recentering. In Table 6, in the left-hand column, is a list of ten such factors, beginning with those which are probably the most effective. (The more precise explanation and analysis of these factors will follow in the next paragraph.) The six problems are given side by side above, ordered from left to right according to diminishing quotients $\frac{S(w.p.)}{S(a.p.)}$. If one of the six problems contains one of the ten hindering factors, this is indicated by a plus sign at the proper place in the table. A positive correlation between the quotients of the different problems on the one hand and the number and probable weight of the hindering factors on the other hand should therefore be noticeable in an increase of plus signs from right to left.

TABLE 6

	Cork	Box	Pliers	Paper-clips	Gimlet	Weight
1. No signalling of the perceptual properties of the crucial object	+	+	–	–	–	–
2. F_1 still quite real	+	+	–	+?	–	+
3. F_1 habitual for the crucial object.	+	+	+	+	+	–
4. The crucial obj. not familiar as "differently applicable"	–	–	–	+	+	–
5. F_2 not familiar as realizable by different objects	+	–	+?	+	+?	–
6. The crucial obejct must first be altered for F_2	+?	+?	–	+	–	–
7. F_1 given really (not merely "in thought")	+	+	+	+	+	–
8. The crucial object individually identical in F_1 and F_2	+	+	+	–	+	+
9. The crucial object not very suitable for F_2	+	–	+?	–	–	–?
10. The crucial object not ready for F_2 as a result of F_1	+	+	+	+	–?	+

(Explanations: + means that the hindering factor is present.
– means that it is not present.
+? means that it is probably or to some degree present.
–? means that it is hardly present.

6. ANALYSIS OF THE FACTORS WHICH HINDER RECENTERING

For the understanding of Table 6, an analysis of the ten hindering factors is required. We shall treat them in order.

1. "No signalling of the perceptual properties of the crucial object." In all six problems, an application or function of the object sought is originally anticipated, "something to . . ." But for a search in the perceptual field—in our problems, the table is searched with the eyes—such a functional and topical signal is too vague, too *unprägnant* (see Chapter VI, 6). Visual search concerns visual properties. *The functional and topical anticipation must therefore be transformed into an anticipation in terms of perceptual content, into a signalling of visual contents, in order to be prägnant, to "hit."*

Example: Something is sought "with which to drive a nail into the wall" (see the weight problem). This topical anticipation forthwith arouses the visual image of a 'hammer' or of an 'object like a hammer, i.e., hard and heavy' (transformation into signal in terms of content). And not until there is such a visual model is the visual search begun.

How promptly the original topical anticipation may lead to an (approximate) anticipation in terms of content follows from these remarks of Ss in the weight problem: "I am accustomed to use as a hammer whatever is at all solid and heavy," or: "Often enough I take a stone, if I have no hammer." It is generally true that, the more typical the function F_2 is for the crucial object or its like, the more easily the original signal by function is transformed into a signal by content appropriate to the crucial object.

Analogously in the pliers problem, the functional and topical anticipation: "a support for the board," immediately calls forth the anticipation by content: "something long, solid. . . ." Just so, the anticipation: "something from which to hang a cord," suggests something in the form of a hook or of an eyelet.

Two of our problems—the cork and the box problem—are in this respect worse off than the rest. Here, as a rule, the original functional and topical signal did not succeed in arousing an adequate model of search in terms of content. It could be seen quite clearly how in these two problems the visual search frequently took place under the original function-signal as such ("something to fix the bar which is a bit too short," or, "something to fasten the candles to the door"). But that means that here—*faute de mieux*—the search is with an *unprägnant* signal.

To such anticipations of function, not defined by content, quite different objects may correspond. The candles could simply be fastened to the door somehow or other with tacks or with the help of a cord or of a plug—solutions which were often actually tried. In other words, the anticipation was not specifically directed to something like boxes. It was interesting to see, both from reports and from observed behavior, that, with the box problem, two of the three successful Ss in the a.p. group arrived at the solution in this way: they started from tacks and looked for a "platform to be fastened to the door with tacks." To these Ss, therefore, the tacks suggested a signal already fairly concrete, which in turn could not fail to suggest immediately the visually represented properties: "light material," "supporting surface"—perhaps, as immediately as in the pliers problem the support-function suggests the visual properties: "something long, stable . . . "

Now, if the general function alone is anticipated, there is a "gap" between signal and object. The filling in of this "gap" has to start from "below," from the object.[133] *And it is really this emer-*

[133] See related material on "suggestion from below" in Chapter I, 11.

gence of the new centering (F_2) from the object itself which is hindered by functional fixedness of this object.

This statement will find corroboration in further experimental results to be cited below. There we shall see that a heterogeneous functional fixedness of the crucial object is unable to resist a sufficiently "pointed" (*prägnant*) property-signal. For the time being, the statement will suffice that the condition: "no signalling of perceptual properties of the crucial object", radically hinders a recentering.

2. "F_1 still quite real" means that at the time in which F_2 becomes real the function F_1 is still itself psychologically real, still "lives" as function. The boxes, e.g., persist in their ("static") function of containers. On the other hand, the ("dynamic") pliers-function of the pliers actually ceases along with its use as pliers. Here, therefore, no more than after-effects of the function F_1 exist at the time when F_2 becomes real. In a more general sense of the word (if we include after-effects in "reality"), F_1 is of course in both cases "still real."

This general factor, the "overlapping of the spheres of reality of F_1 and F_2," will also be subjected to closer examination in the report on further experiments. Here its immediate plausibility may suffice: if something like functional fixedness exists at all, it must be the greater, the more real F_1 still is.

A few examples from other experiments and observations: A child builds a tower. This collapses. A block remaining upright promptly becomes a "soldier," and when the "soldier" falls, it at once becomes a "sword."[134] This chain of recenterings is made possible, *inter alia,* through the fact that every time a structure is destroyed from without (cf. the collapse of the tower), the old function vanishes, so that the object becomes once more relatively *neutral.*—This same reduction of 'reality' can be caused by "satiation." The recentering of play-things to be observed so often with children typically appears after they have played with a given object for some time and after satiation has therefore set in. According to Karsten, satiation finds its clearest expression in striving for variation. On the other hand, often it is indubitably the unfolding of the F_2-situation which destroys the 'reality' of the old situation and of the function F_1 indigenous to it. In reference to this, an observation of my own: I lay the pencil as bookmark between two pages, while I read something at another place in the book. I wish to make a note on what I have read here, and unhesitatingly take for the writing (F_2) the pencil, whose function as a bookmark then naturally becomes illusory. This is facilitated by the further fact that the function "for writing" is the habitual function of the pencil (see page 92 above).

3. "F_1 habitual for the crucial object" means that the function F_1 has really passed into the "flesh and blood" of the crucial object, and can now be called its "quasi-property." Now this is not the case with the weight problem. A weight is for weighing, but it is by no means familiar as pendulum weight. To be sure, a weight may originally have about as much affinity to a pendulum weight (F_1) as to a hammer (F_2). An object is of course especially easy to recenter when F_2 represents its original function which is only temporarily supplanted by F_1. (Example: a large log of firewood, still to be chopped, which has served as chopping block for its like, is itself eventually chopped up.)

[134] According to Muchow; see Scheerer (34), p. 232.

4. "The crucial object not familiar as 'differently applicable'." It is clear that a heterogeneous pre-utilization will "fix" the object the less, the more this object already has the character: 'variously applicable'. A box, a pair of pliers, e.g., are probably less specialized in function than a paper clip or a gimlet. Thus, pliers are often used as substitute for a hammer, a box frequently as support.

A parallel from Köhler's experiments: "Besides, the blanket, is seen and used daily, and is thus unique and in a different category from other objects." (20, p. 38, footnote.) This is given as partial explanation of the fact that the blanket was relatively promptly used as substitute for a stick. The effect of the daily handling can hardly be conceived otherwise than in the sense of our factor of variable applicability—in connection with the factor of "contact".

5. "F_2 not familiar as realizable by different objects." Some functions are "fixed" from the start to quite definite objects; other functions may be realized by rather heterogeneous objects. The statement of a subject (p. 92): "I am accustomed to use anything suitable as a hammer," points directly to the fact that the hammer function does not tend to be very fastidious in the selection of its objects. In the same way, the function F_2 in the box problem: "something on which to put . . . " has, of course, countless possibilities of realization in objects. In the course of time one does put almost everything on about everything else.

6. "The crucial object must first be altered for F_2." This factor is unambiguously present in the paperclip problem. A paperclip which is unbent and a proper paperclip have not much more in common than their material. To be sure, in the box problem, and similarly in the cork problem, the necessary alteration does not happen to the crucial objects *per se;* but it does happen to the visual whole of which the crucial object is a part. An empty box is visually something other than a filled one, an isolated stopper something other than one "sticking in" a bottle. An alteration (in our case, a rupture) of a whole alters the phenomenal character of the part.

7. "F_1 given really (not merely "in thought") means that F_1 was or is an actual "fact", that it is not merely ideally ("merely psychologically") given—as is the function of the weight as a pendulum-weight in the weight problem. (Despite the string fastened to it, the pendulum weight would be fully realized only if the pendulum were hung up.)

8. "The crucial object individually identical in F_1 and F_2." Only in the paperclip problem does F_2 not take place with the identical object of F_1, but merely with a representative of the same *genus proximum.* A whole genus may be functionally "fixed".

9. "The crucial object not very suitable for F_2." This factor is related to Factor 1. The less adequate F_2 is for the object, the more difficult is the recentering into F_2. Pliers and cork were sometimes perceived as not especially suitable for F_2, and once this happened with the weight.

10. "The crucial object not ready for F_2 as a result of F_1". This factor is just about the rule in our problems. Yet it happened in the gimlet problem that the gimlet used for boring (F_1) obtruded itself as a thing on which to hang the cord because it was already sticking in place.

By this discussion, the ten hindering factors in Table 4 ought to have become concrete. Now if we look at Table 4, at least one thing must strike us immediately: the great difference between the first two problems and the sixth, in respect to the number and weight of the hindering factors. We have here all the correlation with our quantitative results which one can wish for. Also in respect to the difference between the first two and the last four problems, the correlation is fairly good—as good as one can expect with factors considered merely in a qualitative, not in a quantitative sense.

7. EXPERIMENTS ON THE 'REALITY' OF F_1

On the question of the 'reality' of the earlier function (F_1), I undertook a series of experiments which would at the same time test the main result in a somewhat modified experimental procedure. I started from the assumption that the functional fixedness of an object would have to be less if F_1 did not belong to the same problem-context as F_2. To test this assumption, problems in three variations were set up (of which only the one named in second place—*b*—is new):

a) F_1 and F_2 belong to the same problem-context.

b) F_1 and F_2 belong to different problem contexts which, however, follow in quick succession and have the same inventory of objects.

c) F_1 is lacking.

Three problems were used (one old and two new).

1. The *pliers problem* (see p. 86 above). The *variation b)* here required a new but extremely simple problem in which the pre-utilization of the pliers had to occur: the S, sitting on a chair, had to get hold of an object far enough away so that a stick was necessary. The stick was nailed to a board (which functioned in the F_2-problem as "flower-stand") and was to be freed with the pliers. Immediately following,[135] the pliers problem as described above (p. 86) was presented, of course, in the w.p. setting—in other words, as in the *c-variation*.

2. The *book problem*.

Variation a: The S was to raise a projected image through lifting the lens of the projector, and the extent of this had to be calculated by reference to a table of logarithms (F_1) in a book—the crucial object. By far the simplest way to lift the lens was then to set it on the book (F_2).

Variation b: Here a series of computations, the last one with logarithms, was to be carried out. Following this, the problem above was then immediately presented—now, of course, without reference to the table of logarithms, i.e., as in the *c-variation*.

3. The *yardstick problem*.

Variation a: Two pendulums of different length were to be hung from the table-top side by side ("to test the dependence of the period on the length of the pendulum"). The length of the pendulum was to be measured (F_1) with a (folding) yardstick (crucial object). As support of the pendulum (F_2) the yardstick was by far the most suitable object if it was screwed to the table top.

Variation b: A problem concerning visual length perception, at the end of which the yardstick was used to test the accuracy of achievement. The problem above, omitting measurement of the pendulum length, followed immediately.

The experimental procedure was similar to that in the preceding experiments. But more objects, more carefully selected, were offered for choice on the table. My purpose was to put the main emphasis on the number of pre-solutions, instead of, as formerly, on the number of solved and unsolved problems. For this reason, the Ss were also urged to produce new proposals tirelessly until they reached the correct solution. No time limit was set in this case, except for that imposed by complete "exhaustion." Unfortunately, however, this method of experiment and

---

[135] Lest the two variations *a*) and *b*) differ in respect to the time-interval between F_1 and F_2, care was taken that each time in the b-variation a clear hiatus, but no pause, occurred between the two part-problems. Therefore, the instructions for both problems were given together at the beginning.

of evaluation proved to be less appropriate than the earlier one. A pre-solution is more difficult to observe than the fact that the problem is not solved and, beyond this, it depends more on chance influences.

Each S was given all three variations—distributed among the three problems—

TABLE 7

	Variation a	Number of Pre-solutions in Variation b	Variation c
Cycle I*	25	13	15
Cycle II	19	7	10
Cycle III	14	12	10
Total	58	32	35

* In cycle I, the *a-variation* was twice not "correctly" solved at all, while in all other experiments the correct solution was eventually found.

and, altogether, the three variations fell with equal frequency on the three problems. In all, we worked with nine Ss: three "cycles" were thus realized. In each cycle, each of the three variations was carried out on three Ss.

The results (the numbers of pre-solutions) from all three cycles are contained in Table 7.

From Table 7 it can be seen:

1. that the separation of the two functions F_1 and F_2 in two independent problem-situations about eliminates the disturbing functional fixedness of the crucial object. The results of a few seminar experiments[136] set up exclusively with the pliers problem were similar (see Table 8). Obviously, the psychological "reality" of F_1 is restricted to the problem situation in which F_1 occurs. At any rate, beyond this area the reality of F_1 is abruptly lowered.

2. From a comparison between *a-variation* and *c-variation* a complete confirmation of our main result follows also: *heterogeneous fixedness, with the old function still 'real', hinders recentering if the F_2-signal is relatively unprägnant.*

8. EXPERIMENTS WITH A MORE PRÄGNANT SIGNAL

The thesis has already been formulated and supported that heterogeneous functional fixedness of an object hinders solution especially if the signal is relasively *unprägnant.*[137] Since this statement has great significance for the theory of the "finding of a solution through resonance," I have endeavored to confirm it by experiments.

A very simple observation: Suppose that, in the box problem, one or all of the boxes are green, while all other objects are not green, or very little so. If the following instruction is now given:

TABLE 8

S	No. of Pre-solutions in Variation a	S	No. of Pre-solutions in Variation b
1	2.5*	5	0
2	1	6	2.5
3	4	7	1
4	6.5	8	3
Arith. Mean	3.5		1.6

* Those pre-solutions which were either not taken quite seriously by the S (e.g., which had wish-character) or which represented renewal of earlier proposals were valued as "half" pre-solutions.

[136] Here individual differences could not be eliminated. Four Ss were given the a-variation and four other Ss the b-variation.

[137] We found in Chap. VI that "*unprägnant*" by no means coincides with "ambiguous". The most definite, least equivocal signal can be completely *unprägnant*, i.e., it can unambiguously correspond to the object without hitting its characteristic nature.

"The solution-object is green. Look for something green," the box will be found at once, despite ever so 'real' fixedness. The same thing happens if some other visual property of the crucial object is anticipated by the signal. Here hardly any trace of a difficulty in recentering will be found.

An explanation for this was already given in an earlier passage (see p. 93 above). Only if the signal is *unprägnant,* or does not reach into the autochthonous nature of the crucial object, if therefore there is still an unbridged distance between the content of the crucial object and that of the signal, and if consequently the crucial object (like any neutral object) must be "evaluated" for its possible suitability—only then does heterogeneous functional fixedness have a hindering effect. It is the development of the new centering from the object itself which is hindered by the fixedness. The central significance of this fact was already discussed in Chapter III, 4, under the title: *A More Precise Demand Facilitates Restructuration.* See also Chapter VIII, 8.

Whereas in the observation just mentioned the signal was extremely *prägnant,* one that is less *prägnant* is involved in the following experiment. The box problem was so altered that only *one* candle was to be fastened to the door, and that one box filled with candles was on the table among other objects. One group of 4 Ss—the P-group—was given the problem with a relatively *prägnant* signal, 4 other Ss—the U-group—with a relatively *unprägnant* signal. After the general problem had already been described in other respects, the P-group received the instruction: "Use for the solution the tacks and something which can easily be fastened to the door with tacks." The U-group, on whose table there were no tacks, received the instruction: "Not everything which you need for the solution is on the table. You may ask for what is lacking."—While in the first case, therefore, a signal by function was given, through which the decisive properties of the object (platform, light material) were very strongly suggested, there was in the second case only the noncommittal anticipation: "something for fastening the candle to the door." Even the tacks, which could easily have suggested a *prägnant* signal, had been removed from the visual field of the Ss. —A third group of 8 Ss,[138] the I-group (intermediate group), received the problem in the "normal form," i.e., tacks were on the table. Only the anticipation contained in the general setting of the problem: "something for fastening the candle to the door," was here explicitly given. But from the tacks there issued at least a suggestion of the signal (which was explicitly added for the P-group)—hence the name: "intermediate" group.

Table 9 contains the results of this experiment. (From the statistical point of view, I admit, these results are not yet quite conclusive.)

Explanations: the results were evaluated according to the number of pre-solutions and to the total time up to the final solution (thus according to two parameters dependent on each other). The pre-solutions were divided into three classes: whole, half and quarter pre-solutions. Thus instead of the number of pre-solutions, the "measure of pre-solutions" was used. What counted as a "half" pre-solution, has been explained in the footnote to Table 8. I counted a "quarter" pre-solution if an object was only "grazed," i.e., touched or picked up quite briefly. Since this ordering to quite definite fractions of

[138] The double number of Ss is explained by the fact that originally two groups were here planned: for one, the tacks were put right next to the box; for the other, at a greater distance. This difference, however, proved to be of no importance.

TABLE 9

P-group			I-group			U-group		
S	Measure of pre-solutions (see below)	Time (sec)	S	Measure of pre-solutions	Time (sec)	S	Measure of pre-solutions	Time (sec)
1	0.75	40	5	1.5	30	13	3.0	105
2	0	0	6	1.0	15	14	5.75	150
3	0.5	5	7	3.5	65	15	2.0	25
4	1.25	40	8	0	0	16	0	0
			9	0	0			
			10	0.5	8			
			11	1.0	10			
			12	2.0	70			
Arith. Mean	0.6 (1.5)	21		1.2 (1.5)	25		2.7 (3.2)	70

unity contains arbitrary elements (which are, however, uniformly distributed) I add in brackets, as a further control, the arithmetic mean of the "number of pre-solutions", i.e., of the pre-solutions counted as integers.

Table 9 shows that the more *prägnant* the signal (or the more suggestive a *prägnant* signal), the easier the recentering, i.e., the lower the measure of presolutions and the total time. The times and the "numbers of pre-solutions" do not show this correlation between P-group and I-group—yet, they do not show a negative correlation.

9. PRÄGNANT SIGNAL AND FIXEDNESS RELEVANT TO THE SITUATION

Suppose that in the box problem, where this time only one candle is to be fastened, two boxes are put among other objects on the table, one box filled with material central to the situation (candles) and one with material peripheral to the situation (buttons). Suppose, moreover, that the following two premises are valid: (1) the less *prägnant* the signal, the more effective a fixedness; (2) an object charged with a function relevant to the situation is more fixed than one charged with an irrelevant, peripheral function—while the former has more "contact" with the subject than the latter. If this be granted, then under the above experimental circumstances, the chances of the box of buttons must increase, the less *prägnant* is the signal. Thus the experiments just indicated were carried out. One group of Ss was given the *prägnant* signal (introduced in 8), another the usual, *unprägnant* signal. Table 10 shows the result.

TABLE 10

	P-group (12 Ss)	U-group (14 Ss)
candle box	chosen* 10 times	chosen* 7 times†
button box	chosen once	chosen 7 times
both at once	chosen once	chosen 0 times

* Or merely: "The thought of using a box came to me in connection with this box."
† With two of these 7 Ss it could clearly be followed how *spontaneous* formation of the prägnant anticipation preceded the choice of the candle box. The two Ss sought expressly for something which could be fastened with nails (which they requested).

Table 10 shows that the box of buttons

actually increases its chances if the signal is less *prägnant,* and if consequently the greater fixedness of the box of candles becomes still more decisive.—If Ss who had chosen the box of candles or that of buttons were asked why the thought of using a box came to them in connection with this and not with the other box, the answer was frequently: "because it had something to do with the candles," or: "because the candles were in the other." Of course this could also be rationalization.

10. ON HOMOGENEOUS FIXEDNESS AND ON TRANSFER

Until now, only heterogeneous fixedness has been discussed. What really happens in the case of *homogeneous fixedness?*

An extreme case: Let F_1 be "something on which to hang a whisk broom"; let F_2 be "something on which to hang a shoehorn". Will F_1 disturb F_2? No, on the contrary. F_2 will remind of F_1, and the use of the crucial object in the function F_2 will therefore *profit* from its previous use in the function F_1. In other words, a *transfer* will take place.

To use an example cited by Selz for "means-abstraction conditioned by chance:" When a soaring kite brought Franklin to the idea of the lightning-rod, the kite's function F_1 (specifically the "reaching up high") suggested the function F_2 ("some such thing could catch the lightning"); it certainly did not suppress it. Kite and lightning-rod both contain the same functional components, just the "reaching up high."

The ground for such a facilitation of the solution may be very precisely formulated: *The fact that F_1 and F_2 are homogeneous means simply that by F_1 the crucial object is expressly "tuned" to F_2.* But this again means that the F_2-signal becomes *prägnant* in respect to the object "prepared" by F_1. Thus homogeneous fixedness has as consequence that the F_2-signal becomes particularly *prägnant.* Hence the solution is facilitated.

For transfer to be possible, F_1 and F_2 need not be just the same function. If one considers the functions concretely enough, they are always different. "Hook for the whiskroom" is not the same as "hook for the shoehorn". But what matters is merely that the two functions be based on the same general function in the same way, ultimately on the same properties (dispositions) of the crucial object.

Each function of an object is based on certain properties of this object, for instance, the function "ruler" of a bar is based on the straightness of its edge, the function "stick (to lengthen the arm) is based on length. Or—to remain with our crucial objects—the properties: "levers, crossing each other, firmness . . ." make the pliers into "pliers", the properties: "length, firmness . . ." into "support".

An example for two quite different functions which are nevertheless founded on the same properties of the object: F_1, "something from which to hang two pendulums side by side"; F_2, "something over which to lay a flower stand". Both functions are based on the properties: "lengthy", "stable".

In all cases of functional fixedness which played a part in our experiments, F_1 and F_2 were based on *different* properties of the crucial object, more correctly: on two different "reliefs" of the properties. (For it does not matter whether a property merely enters into the function concerned; rather it matters how, at what place, and with what weight it enters in.) Moreover, there was not any more basic and general function which F_1 and F_2 had in common. In short, there was not the slightest cause, no foundation whatever, for a recognition, no reason why the object used in F_1 *should* "resonate" when the signal F_2 was given.

11. THE DYNAMIC MEANING OF HETEROGENEOUS FUNCTIONAL FIXEDNESS

It is now time to raise the question: *What sort of alteration does an object undergo through heterogeneous functional fixedness?*

As I see it, there are three kinds of alteration:

1. By means of F_1, the crucial object is embedded in a particular context, in a functional whole, which is to some degree dynamically segregated. In this fashion the object is *"absorbed"*, *"capsulated"*. If this functional whole disintegrates—see the factor of "reality" (7 above)—its parts, the functions, certainly die with it, but naturally not in its elements, the objects. These it "releases from its grasp".

2. Through F_1, *the relief of properties in the crucial object is altered.* The properties particularly claimed by the function F_1 stand out, become dominant, "central" (hence the expression: "recentering"). Those claimed less or not at all recede, and sometimes drop out completely.[139] The crucial object is so to speak specifically "polarized" by the forces of the functional "field."

3. In the degree to which F_1 and F_2 belong to the same comprehensive whole and are experienced as mutually required functions of this whole, a curious factor enters into play. The crucial object is expected to change over from one function into *another* function of the *same* comprehensive whole, i.e., into a function which is in active relation with the first. And this *"shift of function within a system"* frequently offers considerable difficulties to thinking.

[139] See Chap. VI, 1, on the recentering of an object or field as to content when it is under a specific pressure of search.

As clarification of the concept of "shift of function within a system:" I call two different functions of the same object "contrary". This is a generalization of the logical concept: "contrary terms". For example. one calls long and short (or red and blue) contrary, because in a pure form they are mutually exclusive. They belong to the same "dimension," concern the object in the same respect, have the same structural locus, and consequently are in specific and active relation to one another. Long and red, on the other hand, are not contrary terms; between them there is, in a way, a "dead interval."

Until now the concept "contrary" has been defined in logic only with reference to abstract or ideal wholes (namely, property dimensions). I apply it analogously in relation to "real wholes," i.e., to particular and sometimes unique real structures in which different functions demand each other in different places, such as, e.g., hammer and anvil, father and son, radius and tangent. Such functions may be called "really contrary." Now, if one and the same object is to take on in succession "really contrary" functions, we shall call this a "shift of function within a system."

All three factors probably play a rôle in heterogeneous functional fixedness as it appears in our experiments. The third factor is probably the least important, since our problem situations are not very "strong" (functional) *gestalten* (W. Köhler), and since consequently the intra-systemic functions F_1 and F_2 are in only weak "contrary" relation if in any.[140]

It is certainly probable that both factors 1 and 2 are here effective, yet it must be pointed out that in actual fact all our results could be explained by either of the factors alone. After all, we do not know their quantitative potency.

On the other hand, it may at first appear as though several arguments could be raised against an influence of factors 1 and 2 in our experiments. Against 1, one could object as follows: With homogeneous fixedness, which involves "capsulation" (factor 1) but no alteration of the relief of properties (factor 2), no inhibition, no hindering of the solution takes place, but on the contrary a facilitation. On this account, "capsulation" cannot decisively hinder the solution. This objection sounds very plausible at first. In reality, however, it is not valid. The "homogeneity" of the fixedness means (as was already observed) that when F_1 has occurred just the function (or property) F_2 of the object is actualised, brought into readiness, or even made a quasi-property of the crucial object. But through this the signal becomes *prägnant* in respect to the object. Now, we know that a *prägnant* signal may overcome fixedness. Thus that objection to the factor of capsulation becomes untenable. Homogeneous fixedness does mean capsulation; but it means at the same time that the signal becomes more precise, and this enables the search to penetrate all "capsule walls."

Against factor 2 as well an objection suggests itself at first. Suppose that in heterogeneous fixedness the relief of properties in the crucial object were really and essentially altered, so that the parts of the relief corresponding to the function F_1 became dominant, and those corresponding to the heterogeneous function F_2 became in contrast recessive. Then, with heterogeneous fixedness, how could an F_2-signal ever be *prägnant?* For, those properties of the crucial object which correspond to F_2 would always be relatively recessive. To this I should like to answer: 1) With heterogeneous fixedness a maximally *prägnant* signal may not in fact be possible. 2) The heterogeneous deformation of the relief of properties is probably not so great that a search directed to certain properties could not find these properties despite the deformation.

---

[140] From the finding that the fixedness virtually disappears if F_1 and F_2 belong to expressly different problem situations, one could be tempted to infer that the third factor, too, has considerable influence. However, that finding can be fully explained by simple loss of "reality" (disintegration) of the functional whole to which F_1 belongs, therefore by exclusive application of the factors 1 and 2.

CHAPTER VIII

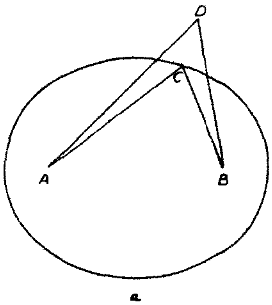
ON FUNCTIONAL FIXEDNESS OF MATHEMATICAL SOLUTION-ELEMENTS

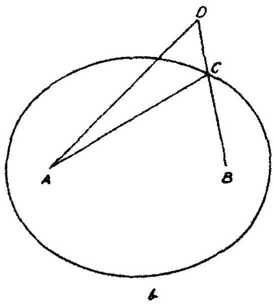
(A Contribution to the Problem of "Mathematical Talent")

1. ON RESTRUCTURATION OF MATHEMATICAL SITUATIONS

In Chapter III, which dealt specifically with mathematical solution-processes, the conjecture was made that the "many-sidedness" and restructurability of a (subjective) thought-material depend not alone on familiarity with the pertinent general relationships (theorems) and aspects, but also and quite essentially on certain subjective constants of the material. In order to test and, if possible, to corroborate this conjecture, we shall now investigate what kinds of "restructurations" thinking has to perform in mathematics. With this we touch upon a problem which is very closely related to that of the previous chapter.

There are certain crucial points in thought-processes at which individual differences are sharply shown up. We have already had opportunity to become aware of such points. We remember, for instance, the restructuration of a six-place number of the type abcabc into abc thousands plus abc units. But there are many more characteristic examples. Let us consider the very simple task of showing that all points outside an ellipse are at a greater distance from the two foci than are the points on the ellipse, for instance, AD + BD > AC + BC. (See Fig. 23a.) After we have substituted for C an ellipse-point on AD or BD (see Fig. 23b)—the equivalence of all ellipse-points is already assumed—the proof requires a restructuration which is simple in itself, but (somewhat paradoxically expressed) extraordinarily difficult in comparison with its simplicity. Instead of conceiving the lines as they were introduced—homologous to each other—namely, AC with AD and BC with BD, for the purpose of proof one must consider AC in relation to AD + DC and BC in relation to itself. Then what is to be proven follows immediately; for, AC, as straight line connecting A and C, is shorter than AD + DC, and BC is equal to itself. The applicability of the decisive proposition to the given facts thus presupposes a radical *restructuration* of the definitely structured situation. By the manner in which it was introduced, the given situation was *fixed* in a definite way, heterogeneous to the solution. There are people to whom such restructuring comes easily, and others to whom





FIGS. 23a and 23b

it comes with difficulty. And yet the proposition that the straight line is the shortest connection between two points may be equally familiar to them all.

Another example is offered by the solution-phase: "$a.b$ = common multiple of a and p" in the process cited above on p. 40. The numbers a and b were introduced in the same rôle in respect to the prime number p. Now $a.b$ is suddenly to be conceived as common multiple of a and p; i.e., the numbers a and p, which were thus far in a certain contrast to each other, are now suddenly to be treated as having the same rôle. Of course it is evident to everyone that this can be done; but the spontaneous conception is made very difficult to some people by the necessary restructuration. One more example: For many Ss it is difficult to comprehend spontaneously that $1 \times 2 \times 3 \times \ldots \times 1000$ represents a multiple of each of the numbers from 1 to 1000. 5×7 is for them a multiple of 7, but not—simultaneously—a multiple of 5 as well.

For several years I systematically investigated one subject as to difficulties in thinking. The unambiguous result was that "onesidedness" of (mathematical) thought-material was to a high degree characteristic for him, in contrast to mathematically more gifted persons, and could therefore with full right be regarded as a constant of his intellectual makeup.

It is not yet explained with what sort of a "structuration" we are here dealing, what sort of claims are raised that in many cases thinking finds so difficult to satisfy. But first, let me make some preparatory remarks.

2. THE FINDING OF A PROOF AS RECOGNITION INVOLVING ABSTRACTION

Mathematics, as a strictly deductive science, tries to reduce all its propositions to as small as possible a number of axioms. Each proposition, be it ever so directly evident to inspection, must legitimate itself as deducible from the axioms (provided that it was not itself chosen as an axiom).

The applicability of a few axioms to numberless concrete facts requires *abstraction from certain perceptual properties* of the "models"[141] concerned. One must abstract from them in order to *recognize* an axiom in a concrete present fact. For, frequently the perceptual properties of the axiom-model and those of the case of application are in a certain *contrast* which hinders recognition.

Thinking has something to achieve in the abstraction from these perceptual properties, which are psychologically real but irrelevant for the deductive context, and this achievement appears to be difficult in very different degrees for different people.

What is the character of these crucial perceptual properties? And can differences in the ability to abstract from these conflicting perceptual properties, when one fact is to be recognized in another, really be regarded as a decisive factor of mathematical thinking?

3. ILLUSTRATION OF THE PERCEPTUAL PROPERTIES OF MATHEMATICAL OBJECTS WHICH HINDER RECOGNITION

For the illustration of those crucial perceptual properties and their incompatibilities, I have chosen three elementary mathematical proofs, in which what is to be proven is demonstrated with the help of one and the same axiom. Let it be expressly emphasized that the three examples are not exceptional in mathe-

[141] By a "model" we understand the perceptual or ideational substratum of the concrete fact concerned, or of the axiom in question.

matics as to the features which will be demonstrated in them; soon this will also be obvious to a reader who surveys his own mathematical experience.

At the bottom of all three examples is one and the same axiom, the axiom of M. Pasch (briefly: P-axiom), which expresses the following directly evident fact: Given a triangle and in the same plane a straight line *which passes through no apex of the triangle.* It follows that, "*if the straight line intersects one side of the triangle, then it intersects still another side of the triangle.*" (See the "model" of the P-axiom—briefly: "P-model"—in Fig. 24.)

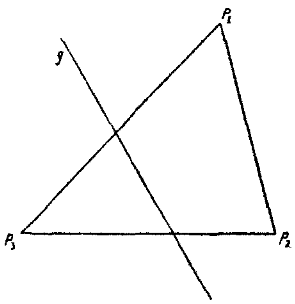


Fig. 24

In the first of our examples, not the P-axiom itself is used, but rather a proposition following immediately from it, namely the proposition: "If the straight line intersects one side of the triangle, then it intersects *one and only one* other side of the triangle." This statement, whose proof with the help of the P-axiom will constitute our third example, we shall call the "P-proposition", to distinguish it from the P-axiom. Along with a few simple axioms (e.g., that two straight lines intersect only once) and many derived statements, this P-proposition is for the moment supposed to be familiar to the subject. One day he finds before him the problem of proving the following proposition:

Premises: Given a straight line g, on g a point A, and through A a second straight line b, different from g, on b two points P_1 and P_2 such that A lies between P_1 and P_2.

Proposition: Then any arbitrarily chosen third point P_3 of the plane may be connected either with P_1 or with P_2 by a straight line in such a way that the connecting line does not intersect the straight line g. (This proposition implies that each straight line divides the plane into exactly two "provinces;" see Fig. 25.)

The proposition can in fact be proven directly by means of the P-proposition:[142] The three connecting lines P_1P_2, P_1P_3, and P_2P_3 may be conceived as the three sides of a triangle of Pasch (briefly, of the P-triangle), and g as the straight line of Pasch (called "transversal").—Then the P-proposition means simply that the straight line g, which by premise intersects P_2P_1, must intersect one and only one other side of the triangle, P_1P_3 or P_2P_3. But this in turn means that P_3 may be connected either with P_1 or with P_2 in such a way that the connecting line does not intersect the straight line g—*quod erat demonstrandum.*

Suppose one does not know beforehand that the proof should be given by

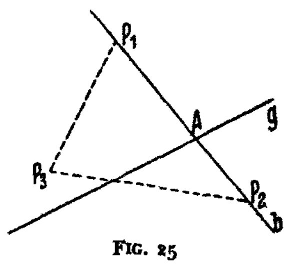


Fig. 25

means of the P-proposition. Then only a series of axioms and theorems are known, but, which axioms are to be applied to

142 Here we are neglecting the two banal cases, that P_3 lies on g or on b.

the given case can be derived solely from the inner properties of the present case, i.e., from the assumptions and the proposition, the goal. Under such conditions, the proof does in fact make certain demands[143] on thinking. For, the perceptual properties ("functions") are opposed to each other in the following characteristic way.

1. The perceptual functions of the elements of *the P-models:*

Here the triangle $P_1P_2P_3$ is given as "primary" and as "absolute", and through this already given triangle, the straight line (transversal)[144] is drawn as "secondary" and "relative" to it. The three triangle-sides P_1P_2, P_2P_3, and P_3P_1 are given "together" and completely "homogeneously" (i.e., in the same function).

2. The perceptual functions of the corresponding elements of the *case to which the P-proposition is applied:*

The straight line g is here "primary", "absolute". Then through a point of this straight line, thus "relative" to it, a second straight line is drawn, and on this the distance P_1P_2 is marked off. Finally and relative to P_1P_2, the point P_3 is placed, and only now do the sides P_1P_3 and P_2P_3 arise as "connections" of P_3 with P_1 and P_2. Thus the three sides of the triangle are here given entirely "inhomogeneously", and the one, P_1P_2, appears more intimately bound up with g than with the other two sides.

It is clear that the crucial perceptual properties are rooted deep in the way the two sets of facts are given to primary inspection, i.e., in the way in which on one hand the constituent parts of the P-model, on the other hand those of the case of application, are perceptually "introduced." More particularly, these perceptual properties are *"functions"* in a very specific sense, i.e., the properties belong to the elements only inasmuch as these elements are at the time parts of one structure or another. Further and above all, it is clear that *the functions contradict each other (are 'contrary').*[145]

1. What is given as "primary" and "absolute" in the case of application, is given as "secondary" and "relative" in the P-model; and vice versa.

2. Some of the facts which are given "inhomogeneously" in the case of application (P_1P_2 inhomogeneous with regard to P_1P_3 and P_2P_3) are given "homogeneously" in the P-model.

3. Some of the facts which are linked in the application (P_1P_2 and g) are not so linked in the P-model. Instead, P_1P_2, P_2P_3 and P_3P_1 are here linked together.

Thus we are now dealing with a particular form of heterogeneous functional *fixedness of the thought-material.* (See Chap. VII.) The structure of the facts concretely present, of the case of application, must be *altered* in characteristic manner if the proposition is to be applied. More particularly, this restructuration is of the kind which we called above (p. 100): "shift of function within a system." The elements are expected to change over from one function not only into some *other* one, but into a *contrary* function. In the present case, these functions are *unspecific,* are functions of formal *Zueinander*[146] as distinguished from

[143] Let the reader not forget that—here and in the following two examples—he does not experience these difficulties directly, since he is expressly told that the proof is each time to be derived by means of the P-axiom (or by the P-proposition). Therefore he approaches the concrete facts of each case with a very *prägnant* direction.

[144] In the term "transversal", this secondary and relative function of the straight line attains *prägnant* verbal expression.

[145] See above, Chapter VII, p. 100.

[146] K. Gottschaldt (7) observed that functions of this same type had a "camouflaging" effect. In other words, they prevented recognition of

the "specific" functions treated in Chapter VII and below, p. 108.

4. CONTINUATION (SECOND EXAMPLE)

We now turn to a second, more impressive example.

Premises: Given two straight lines g_1 and g_2, intersecting at M; further, among other points on g_1, two points A and B so given that B lies between A and M.

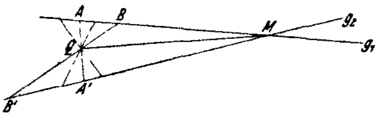


FIG. 26a

Given finally a point Q, between g_1 and g_2. Lines drawn from Q and connecting with A or B or M intersect g_2 at A′ or B′ or M—the projections of A and B and M on g_2.[147] (See Fig. 26a.)

Proposition: Then the "order" of the points on g_1 is other than the order of the corresponding projection points on g_2; i.e., B′, for example, the projection of the point B which lies between A and M, does *not* lie between A′ and M, the projections of A and M.

Who would at first glance recognize the P-axiom in the model of Fig. 26a?[148] And yet it is there. To see it, one need only envisage the two lines MB and MB′ as belonging with BB′, thus with an element which was introduced in connection with and in the same function as AA′ (namely, as one of the many rays which project points on g_1 upon g_2). Then one has a P-triangle. And AA′ must be torn from its community of function with BB′ and the other rays of projection, and be distinguished, singled out, as "transversal." Thereupon one sees: by premise, AA′ intersects the side BB′ between B and B′, but not the side MB (because B should lie between A and M, so that A must lie outside of BM). Consequently, according to the P-axiom AA′ must intersect the side B′M between B′ and M. But this means that A′ lies between B′ and M, and that *therefore B′ does not lie between A′ and M*—which was to be proved.

Such hindrances of recognition in mathematical thought-processes are

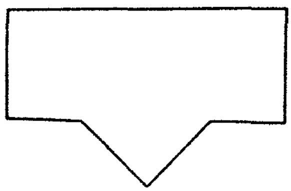


FIG. 26b

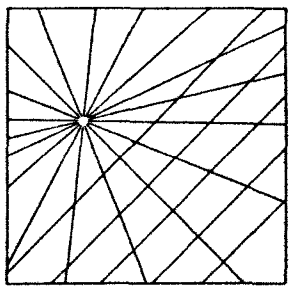


FIG. 26c

closely related to hindrances of recognition which were investigated by Gottschaldt (7). This is most impressively shown by a comparison of our Fig. 26a with Fig. 16b in Gottschaldt's work (see our Fig. 26c). Here the smaller figure reproduced in Fig. 26b is to be recognized, to be *"herausgesehen."*

part-figures in larger pattern, although the former had been made extremely familiar.

[147] That M is its own projection represents one of those "limiting cases" of a concept (here of the concept "projecting point") which are frequent in mathematics. They often go very much against the grain of any thinking that is too much centered on perception or concrete ideation.

[148] As a matter of fact, the P-axiom is here very much more hidden than in our first example.

5. CONTINUATION (THIRD EXAMPLE)

In the third and last example we shall demonstrate a new and psychologically very important form of functional fixedness of thought-material, or—otherwise expressed—of "contrariness" of "functions" in inspected figures. Our example is the proof of the proposition (which we have called the "P-proposition") that a straight line (transversal) never intersects all three sides of a P-triangle.

The proof of this statement requires an introductory comment which was not necessary to the understanding of the previous examples, but which here becomes relevant. If the modern mathematician axiomatically constructs a mathematical discipline, e.g., geometry (see for example D. Hilbert (12)), he purposely behaves as if he knew not the slightest bit more about the objects concerned than what he has stated about them in the axioms. In other words, he uses the axioms as definitions of his concepts. "Point," "straight line," "plane," "between" . . . are for him not the usual perceptual images, but first pure unknowns, in which nothing is definite and

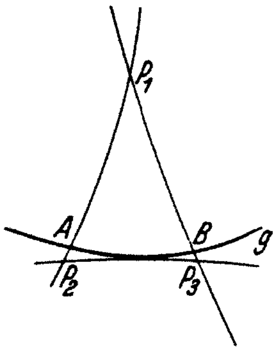


FIG. 27a

known but what follows about them from the axioms in which they appear. Thus, for example, by the axioms which we had to use as premises of our proofs, nothing has yet been established which obliges the straight line actually to appear as 'straight.' We could also, if we chose, imagine some curved line and call it "straight line," provided that this strange straight line fulfills all demands which can be derived from the axioms.

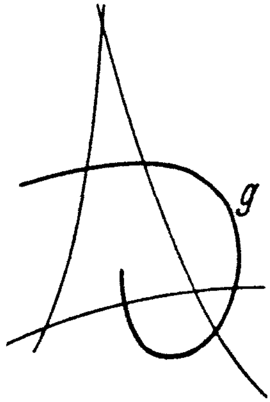


FIG. 27b

From what has been said, Fig. 27a, on which we base our third example, becomes comprehensible.

Let the triangle $P_1P_2P_3$ be the P-triangle, g the transversal. Now, it is to be proved that the fact represented in Fig. 27a is not possible: g cannot come in contact with all three sides of the triangle.[149]

The proof is carried out in a way as surprising as it is elegant, by application of the axiom of Pasch to the special case of Pasch's model which is to be proven impossible. The three lines P_1A, AB and BP_1 may be conceived as the P-triangle, and the line P_2P_3 as the transversal. With this, the proposition that the (former) transversal g can come in contact with all three sides of the (former) P-model becomes equivalent to

[149] The other way in which g could try to intersect all three sides of the triangle (see Fig. 26b) is already barred by the axiom: "two straight lines intersect only once".

the assertion that the (new) transversal P_2P_3 can intersect one side of the (new) P-triangle (P_1AB), namely AB, without intersecting any other side of the same triangle. But this the P-axiom forbids. Therefore it also forbids the proposition that a transversal can come in contact with all three sides of a P-triangle. Q.E.D.

This proof is especially interesting and quite different from the proofs given earlier. There, a perceptual "contrariness" between axiom and special case existed only in respect to such *unspecific* ("formal") functions as "belongingness" and "segregation," "homogeneity" and "inhomogeneity," being given "absolutely" and "relatively."[150] Now, however, there is over and above this a perceptual "contrariness" in respect to quite *specific* ("material") functions as well.

The P-model contains two such specific functions: the function "side of a triangle" and the function which is contrary to this, "transversal." And just these functions undergo shifting: what was "transversal" becomes "side of a triangle," and a "side of a triangle" becomes "transversal." Observation of mathematical thought-processes for years has shown me that this shift of function within a system, which consists in the transformation of specific or "material" functions, makes serious demands on thinking.

[150] Of course, there is perceptual "contrariness" as to such unspecific functions also in the new example. While, in the concrete application, the lines P_1P_2, P_2P_3 and P_3P_1 are introduced as (1) absolute, (2) belonging, (3) homogeneous, and on the other hand the straight line g and therewith the line AB as (1) relative, (2) segregated, (3) inhomogeneous (singular)—in the use of the P-axiom, the corresponding elements are structured quite otherwise. Here the line P_2P_3 is no longer belonging and homogeneous with P_1P_2 and P_3P_1, no longer given as absolute with these. Rather, the lines P_1P_2 and P_3P_1 (more precisely: parts of these lines) have now joined the previously "singular" line AB.

6. SUMMARY

It is the shift of function of the components of a complex mathematical pattern—a shift which must so often occur if a certain structure is to be recognized in a given pattern—it is this *restructuration,* more precisely: this transformation of function within a system, which causes more or less difficulty for thinking, as one individual or another tries to find a mathematical proof.

If a situation is introduced in a certain perceptual structuration, and if this structure is still "real" or "alive," thinking achieves a contrary structuration only against the resistance of the former structure. The degree of this difficulty varies among individuals.

Incidentally, this resistance evinces itself not simply as an obstacle to "finding" but—much more malevolently—as an obstacle to "surveying", more generally: to actual "realization". The points where, during complicated mathematical deliberations, suddenly "the thread breaks," are typically points where one and the same element is suddenly claimed in a contrary function. Contrary functions easily "interfere" with and extinguish one another. Such difficulties in the actual realization of changed structures cannot be discussed here.

7. THE HEURISTIC FUNCTION OF INSPECTION

The summarizing formulation in the preceding paragraph very much needs supplementation. By no means should it give the impression that inspection *per se* is an obstacle to mathematical thinking. In a very general sense, every case of productive thinking relies on inspection. (See p. 52 on inspection as the intermediary for genuine insight.) Inspection may not have the last word, but at any rate it has the first. Its function is essentially heuristic. We remember forms of situation-analysis which belong here: simple reading-off from the model (from the figure embodying the premises), or: "The model reminds of. . . ." Many a property of the premises which is relevant to the proof or to the construction is discovered in inspection much sooner than it is logically deduced. Extremely often, inspection gives direction to logical explication. ("If what looks as though . . . were really valid, then one could proceed so and so. Let me just see if it is valid.") Inspection owes this heuristic function to its "abbreviating" rôle, i.e., to the fact that in inspection, in the examination of given territory as a whole, many implications of the premises are simultaneously surveyable which could become logically explicit only in a clumsy series of successive steps. In this way inspection enables the subject to see the decisive points of attack and to concentrate on these.[151]

Incidentally, it would be a gross mistake to believe that inspection occurs only in geometry. Perceptual models—schemata—play a rôle in all regions of mathematical thinking. It makes no difference that in general they cannot be represented in "drawings." They are mostly much more flexible, more fleeting, more simply indicative than the drawing of a triangle or the like.[152]

On the other hand, what does impede mathematical thinking, specifically restructurations, is not inspection *per se*, but a quite particular way in which it may appear and function. (See p. 110.)

8. RESTRUCTURATION AND PRÄGNANZ OF DEMAND

In Chapter III, a whole paragraph (4) was devoted to the theme: "A more precise demand facilitates restructuration." There, we arrived at the important statement: "A restructuration is the more easily carried out, the more sharply the new structure is envisioned." Similarly we found in Chapter VII (6 and 8) that against a sufficiently "pointed," *"prägnant"* property-signal a heterogeneous functional fixedness was not able to hold out. And again, the restructurations required in the problems discussed in 3-5 of this chapter, can easily be accomplished *if the direction toward the model of Pasch is expressly given.* The most radical shifts of function within a system are often carried out without any difficulty, if one already knows what *proposition* is to be employed and therefore what sort of structure is to be sought in the given material. Here again the heterogeneous structuration opposes only a relatively spontaneous emergence of the new structure out of the given material. With the last-mentioned problems, the original problem-setting may already contain rather specific suggestions, at least for persons of average mathematical talent. Every time, it is a question of discovering relations of "between", of "intersections within and without." Every time, at least four straight intersecting lines are given (or suggested by points). The search for the proof can therefore be under the direction: "What sort of propositions are there which (under such conditions) assert something about 'between', about 'within and without' of points on

[151] It is this very general concept of "inspection" which H. Poincaré has in mind when he says: "We need a faculty which makes us see the end from afar, and intuition is this faculty" (30, p. 218). "Intuition is the instrument of invention" (ibid., p. 219).

[152] For Poincaré as well (see footnote 151), "intuition" is not a monopoly of geometry; *cf.*, for example, the expression: ". . . to perceive at a glance the general plan of a logical edifice." 30, p. 221)

lines?"[153] Since Pasch's axiom was assumed to be familiar *per se,* it must be rather easily attainable by such a signal. In this manner, therefore, a specific direction can arise for the recognition of that structure in the given material which corresponds to the P-axiom. With this direction any contrary structurings of the material can then be overcome without difficulty.

9. DISCUSSION OF TWO HYPOTHESES CONCERNING THE PSYCHOLOGICAL CAUSE OF THE INABILITY TO RESTRUCTURE

Now at last we have come far enough to be able fruitfully to pose the question of the "individual constants" in a given province of thought. I see two different possibilities for explaining why, for many people—for the "poor" mathematicians[154]—the restructurations necessary in finding proofs and constructions seem so much more difficult than for others.

1. The "poor" mathematician is not able to restructure so easily, because his *thought-material is relatively inelastic, rigid, and therefore not sufficiently plastic to be reshaped.*[155]

2. With "poor" mathematicians, the *thought-material is from the very beginning more thoroughly imbued with perceptual functions of the sort described.* For the "good" mathematician, on the other hand, there remains a more abstract stratum into which all those functions, with their "contrariness" to the recognition of other functions, do not reach; in other words, a stratum in which only the specific mathematical properties still exist. In this more abstract stratum, the structures of geometry, for example, consist only in so-and-so-many points, straight lines, planes, and certain relations of incidence and "between" among these elements. The incidence of a point and of a straight line, e.g., is here—quite apart from structure—nothing but incidence, neither a "lying of the point on the straight line" nor a "passing of the straight line through the point." For, again, these two perceptual structurings would already be contrary to each other as to the functions "absolute" and "relative."

We must now take a position in regard to these two hypotheses. The first sounds plausible, but could hardly be valid. For, it may be observed everywhere that from the start, with "poor" mathematicians, those perceptual structurings play in fact a far greater rôle than with "good" mathematicians. There are people, for example, for whom the P-axiom is so much a matter of a triangle crossed by a straight line that they do not grasp the *more general* character of the P-axiom as an axiom of "ordering," of "between." Their "concepts" are too much "image-bound".[156] An especially impressive example: That same S whom I studied for years once learned projective geometry,

[153] In Chap. III (p. 37) we have already become aware of the great significance of this heuristic method which demands "propositions (theorems) about. . . ."

[154] I should like to point out expressly that, according to our findings, one is justified in speaking of "mathematicians" as a general type. Our examples stem not only from geometry but also in part from arithmetic. There exist in fact specific mathematical demands on thinking. Incidentally, this ought to be immediately obvious from the discussion at the beginning of 3 above. Of course, it does not follow that the different mathematical disciplines have no specific demands and ways of thinking of their own. There *is* a specific geometric talent, a talent which is specifically analytical, and one which seems specific to number theory. Every mathematician is aware of this. However, the present investigation will not deal with these specific differences. But let me not therefore be reproached with having ignored them.

[155] By means of a similar theory, K. Lewin has recently attempted to explain certain forms of feeble-mindedness; see (24).

[156] See in this connection Wertheimer, (44).

including the concept of the "projecting figure." By this is meant the connecting of the center of projection with the points and straight lines of a plane in the form of "projecting" straight lines and planes (see the model in Fig. 28a). The S had "understood" it very nicely. But one day he had to recognize the presence of a projecting figure in Fig. 28b (P center of projection, Q any point, PQ the projecting straight line). Quite simply, he felt betrayed. For him a projecting figure had taken on the character of a "pyramid." "Apex" and "base" were consequently entirely non-symmetrical. In short, he did not possess the correct concept at all, but one much more concrete, bound to an image-like model—although just this model was actually meant to illustrate the *generality* of the concept. For this S, it was always difficult to recognize limiting cases as such.

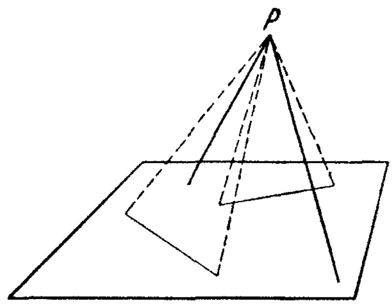


FIG. 28a

Therefore the situation is by no means as though all people accomplished structurations which are, as such, the same, and which function alike, and as though they were in a different position

FIG. 28b

solely—as the first explanation would have it—in respect to the ability to shift from one structuring to another. The second explanation is the correct one: *A different rôle of perceptual structurings, differences as to the degree to which they penetrate the material, these are the causes of the different ability to restructure.*

But how is it that with many people the perceptual structurings dominate so excessively? Clearly, with them these perceptual structures play an indispensable rôle, more or less as do the visual images with people of the visual type. Just as the latter need visual images if anything is to be realized clearly and to be kept stably in mind, *so many people appear unable to make their thought-material precise, to survey it, and to keep it in stable form, unless it is "fortified" and to the core imbued with such perceptual structurings.*

REFERENCES

1. Adams, D. K. Experimental studies of adaptive behavior in cats, *Comp. Psychol. Monog.*, 1929, *6*, No. 1.
2. Claparède, E. La genèse de l'Hypothèse: Étude expérimentelle, *Arch. Psychol.* Geneve, 1933, *24*, 1-155.
3. Duncker, K. A qualitative study of productive thinking. *Ped. Sem.*, 1926, *33*, 642-708.
4. Duncker, K. Lernen und Einsicht im Dienst der Zielerreichung. *Acta psychol.*, Hague, 1935, *1*, 77-82.
5. Einstein, A. *Geometrie und erfahrung*. Berlin: Julius Springer, 1921.
6. Gengerelli, J. A. Studies in abstraction with the white rat. *J. Genet. Psychol.*, 1930, *38*, 171-202.
7. Gottschaldt, K. Über den Einfluss der Erfahrung auf die Wahrnehmung von Figuren. *Psychol. Forsch.*, 1926, *8*, 261-317.
8. Gottschaldt, K. Über den Einfluss der Erfahrung auf die Wahrnehmung von Figuren. II. *Psychol. Forsch.*, 1929, *12*, 1-87.
9. Gottschaldt, K. Der Aufbau des kindlichen Handelns. *Beih. Z. angew Psychol.*, 1933, No. 68.
10. Guillaume, P. and Meyerson, J. Recherches sur l'usage de l'instrument chez les singes. *J. de Psychol.*, *27*, *28*, and *31*.
11. Hertz, M. Beobachtungen an gefangenen Rabenvögeln. *Psychol. Forsch.*, 1926, *8*, 336-397.
12. Hilbert, D. *The Foundations of Geometry*. Trans. by E. J. Townsend. Chicago: Open Court, 1902.
13. Huang, I. Children's explanations of strange phenomena. *Psychol. Forsch.*, 1930, *14*, 63-182.
14. Hume, D. *Enquiries Concerning the Human Understanding* (Sect. 7, Part 2). Ed. by L. A. Selby-Bigge. Oxford, 1902.
15. Husserl, E. *Logische Untersuchungen*. Halle: Niemeyer, 1900.
16. James, W. *The Principles of Psychology*. Holt, 1890. Vol. I, p. 505; Vol. II, p. 346 f.
17. Karsten, A. Psychische Sättigung. *Psychol. Forsch.*, 1928, *10*, 142-254.
18. Katona, G. Eine kleine Anschauungsaufgabe. *Psychol. Forsch.*, 1927, *9*, 159-162.
19. Köhler, W. Komplextheorie und Gestalttheorie. *Psychol. Forsch.*, 1925, *6*, 358-416.
20. Köhler, W. *The mentality of apes*. New York: Harcourt, Brace, 1925.
21. Köhler, W. *Gestalt psychology*. New York: Liveright, 1929.
22. Köhler, W. Das Wesen der Intelligenz. In *Kind und Umwelt*, by A. Keller, 1930.
23. Köhler, W. *Dynamics in psychology*. New York: Liveright, 1940.
24. Lewin, K. A dynamic theory of the feebleminded. In *A dynamic theory of personality*. (Trans. by D. K. Adams and K. E. Zener.) New York: McGraw-Hill, 1935.
25. Lindworsky, J. *Theoretical psychology*. Trans. by Harry R. DeSilva. St. Louis: Herder, 1932.
26. Maier, N. R. F. Reasoning in white rats. *Comp. Psychol. Monog.*, 1929, *6*, No. 3.
27. Maier, N. R. F. Reasoning in humans, I. On direction. *J. Comp. Psychol.*, 1930, *10*, 115-143.
28. Maier, N. R. F. Reasoning in humans, II. The solution of a problem and its appearance in consciousness. *J. Comp. Psychol.*, 1931, *12*, 181-194.
29. Maier, N. R. F. An aspect of human reasoning. *Brit. J. Psychol.*, 1933, *24*, 144-155.
30. Poincaré, H. The Value of Science. In *The Foundations of Science*, trans. by G. B. Halsted. New York: Science Press, 1913.
31. Rademacher, H. and Töplitz, O. *Von Zahlen und Figuren*. Berlin: J. Springer, 1930.
32. Reichenbach, H. *Philosophie der Raum-Zeit-Lehre*. Berlin & Leipzig: W. de Gruyter, 1928.
33. Schapp, W. *Beiträge zur Phänomenologie der Wahrnehmung*. Halle: Niemeyer, 1910.
34. Scheerer, M. *Die Lehre von der Gestalt*. Berlin und Leipzig: W. de Gryter, 1931.
35. Schlick, M. *Allgemeine Erkenntnislehre*. Berlin: J. Springer, 1918.
36. Selz, O. *Ueber die Gesetze des geordneten Denkverlaufs*. Stuttgart: Spemann, 1913.
37. Selz, O. *Zur Psychologie des produktiven Denkens und des Irrtums*. Bonn: F. Cohen, 1922.
38. Shepard, John F., and Fogelsonger, H. M. Studies in association and inhibition. *Psychol. Rev.*, 1913, *20*, 290-311.
39. Thorndike, E. L. *Animal Intelligence*. *Psychol. Monog.*, 1898, *2*, No. 4.
40. Thorndike, E. L. *Human learning*. New York: Appleton-Century, 1931.
41. Tolman, E. C. *Purposive behavior in animals and men*. New York: Appleton-Century, 1932.
42. Tolman, E. C. Gestalt and sign-gestalt. *Psychol. Rev.*, 1933, *40*, 391-411.
43. Wertheimer, M. Untersuchungen zur Lehre von der Gestalt. *Psychol. Forsch.*, 1922, *1*, 47-58; 1923, *4*, 301-350.
44. Wertheimer, M. Über das Denken der Natürvölker. In *Drei Abhandlungen zur Gestalttheorie*. Erlangen: Vevl. d. Philos. Akademie, 1925. Pp. 34.
45. Wertheimer, M. Ueber Schlussprozesse im produktiven Denken. In *Drei Abhandlungen zur Gestalttheorie*. Erlangen: philosophische Akademie, 1925.
46. Yerkes, R. M. The mind of a gorilla. *Genet. Psychol. Monog.*, 1927, *2*, 1-193.

INDEX