

OXFORD



Words and Images

AN ESSAY ON THE ORIGIN OF IDEAS

Christopher Gauker

Words and Images

This page intentionally left blank

Words and Images

AN ESSAY ON THE ORIGIN OF IDEAS

Christopher Gauker

OXFORD
UNIVERSITY PRESS

OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi

Kuala Lumpur Madrid Melbourne Mexico City Nairobi

New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece

Guatemala Hungary Italy Japan Poland Portugal Singapore

South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

© Christopher Gauker 2011

The moral rights of the author have been asserted
Database right Oxford University Press (maker)

First published 2011

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above

You must not circulate this book in any other binding or cover
and you must impose the same condition on any acquirer

British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloging in Publication Data
Data available

Typeset by SPI Publisher Services, Pondicherry, India
Printed in Great Britain
on acid-free paper by
MPG Books Group, Bodmin and King's Lynn

ISBN 978-0-19-959946-2

1 3 5 7 9 10 8 6 4 2

For Alice

This page intentionally left blank

Contents

Preface	xi
Introduction: Defining the Question	1
1. What are concepts?	2
2. Philosophers versus psychologists	10
3. My theses	13
4. Some disclaimers	14
Chapter 1: The Lockean Theory	17
1. The composition theory	19
2. Abstraction-as-subtraction	24
3. Abstraction-as-representation	28
4. Concepts as the building blocks of judgments	31
5. Contemporary Lockeans	33
5.1 A recent version of the composition theory: Rosch	33
5.2 A contemporary version of abstraction-as-subtraction: Mandler	36
5.3 A contemporary version of abstraction-as-representation: Prinz	40
6. A collection of problems	45
Appendix: Locke as trope theorist	46
Chapter 2: The Kantian Theory	50
1. Kant's own theory	50
2. Evidence for this interpretation	53
3. Why we cannot be Kantians today	56

4. Are concepts embedded in perception?	59
5. A Kantian in contemporary psychology: Barsalou	65
6. A Kantian in contemporary philosophy: McDowell	70
6.1 McDowell's Kant interpretation	71
6.2 McDowell on the justification of belief through perception	75
7. The lessons	79
Appendix: Analysis as the representation of synthesis	79
 Chapter 3: Regions of Similarity Space	 86
1. Churchland and Gärdenfors	89
2. Is the theory inevitable?	93
3. An empirical challenge	96
4. Why judgments cannot be built from regions	99
4.1 Learning about individuals	100
4.2 Learning about kinds	102
4.3 Modal distinctions	105
4.4 Learning through language	106
4.5 Complex logical structures	107
5. Thought without boundaries	109
 Chapter 4: The Sellarsian Theory	 112
1. Sellars's functionalism	113
2. The myth of Jones	122
3. Perception	130
4. Right-wing Sellarsianism: Fodor	134
5. Left-wing Sellarsianism: Brandom	138
 Chapter 5: Imagistic Cognition	 145
1. What are mental images?	147
2. Feats of imagination	149
3. The analysis of imagistic cognition	151
3.1 Perceptual similarity space	151
3.2 Object tracking	155

3.3 Acquired dimensions of perceptual similarity space	157
3.4 Imagistic causation	161
4. Animals and babies	163
4.1 A brief review of some of the challenges	164
4.2 Quinn's studies of human infants	171
4.3 Savage-Rumbaugh's studies of chimpanzees	174
 Chapter 6: Similarity without Concepts	 184
1. Some specious objections	185
2. Similarity and commonality	187
3. Imagistic representation	192
3.1 Mere misperception	194
3.2 Persistent illusions	200
3.3 The correspondence between dimensions	203
4. Contra Tversky	206
4.1 The empirical results	207
4.2 The contrast model	209
4.3 Gleitman's better explanation	211
4.4 The remaining results	215
 Chapter 7: Cooperation by Means of Words	 217
1. The paradigm case	221
2. Some further assumptions	224
2.1 Goals	224
2.2 Varieties of imagistic representations	225
2.3 The composition of similarities	227
2.4 Counting demonstratives	228
3. Production and acceptance of literals	231
3.1 Production of atomic sentences	232
3.2 Acceptance of atomic sentences	237
3.3 Negations of atomic sentences	239
3.4 The activation of speech dispositions	242
3.5 An interlocutor's take on the context	243
4. Disjunctions	245
5. General negation	248
6. Psychologic	249

7. Conditionals	250
8. Universal quantifiers	253
Chapter 8: Thinking in Language	256
1. What is inner speech?	257
1.1 Conceptual thoughts and verbal imagery	257
1.2 Consciousness of verbal imagery	259
1.3 Sentences in thought	260
1.4 Beliefs versus occurrent thoughts	263
2. Easy answers to cheap shots	265
3. The utility of intrapersonal discourse	273
4. Semantic norms	277
References	285
Index	299

Preface

When I was seventeen, I rode my bicycle to the B. Dalton bookstore in the Southdale Mall in Edina, Minnesota, and bought the first volume of a Dover reprint of the two-volume Alexander Campbell Fraser edition of Locke's *Essay on Human Understanding*. Lying on my bed reading it, I thought, "Oh! I'd better get the other one", and rode back to get volume two. Soon I was led to Kant. My earliest philosophical ambition was to rewrite *The Critique of Pure Reason*. I thought it could have been clearer. I went to college and considered myself a Davidsonian. I went to graduate school and tried to be a Sellarsian. As I was getting my degree, connectionism came on the scene, and I tried to believe that too. All of these doctrines have failed me. What I have gained nonetheless is a clear conception of what we have all been trying to achieve. An insight of my own, for which I find no particular precedent in the history of philosophy, is that constructive mental imagery—problem-solving by means of imagistic representations—can do much of the work traditionally ascribed to conceptual thought. So what I am attempting in this book is to record the lessons learned and to use that small novelty to open a lot of doors.

I am intensely aware that for each of the topics I discuss, there is much more pertinent literature that I might have read and cited. My excuse for not reading and citing more of it is that books like this, which strive for a synthesis and a new direction, must sometimes be written, and when they are, then, since so much has been written on every topic of conceivable interest, they will have to be written by people like me, who have read deeply in some areas and widely in many areas but not very deeply in every pertinent area. I do recognize, however, that I display some hubris in designating myself for the task. I think I will do a more careful and conscientious job than many others who have appointed themselves to the task and gained more notoriety than I can expect.

In certain disciplines it is conventional to cite everything that has any bearing on one's topic and assemble long bibliographies. Unfortunately, I have not been able to keep track of everything I have read regardless of whether it has had any real impact on my thinking. A benefit to the reader is that he or she can be sure that most of the literature cited here is in some way genuinely interesting. If I do not always cite the literature that the reader

considers most relevant to my themes, I hope the reader will not lord it over me without considering whether my same basic points apply *mutatis mutandis* to that.

Parts of the present text were taken more or less verbatim from articles I have published in journals. Much of chapter 3 and a bit of chapter 6 were taken from my article, "A Critique of the Similarity Space Theory of Concepts", *Mind and Language* 22 (2007), with permission of John Wiley and Sons. Chapter 5, section 3a, as well as figures 7 and 8 of chapter 6 and some of the argument pertaining to those figures, were taken from my article, "On the Evidence for Prelinguistic Concepts", *Theoria: An International Journal of Theory, History and Foundations of Science* 54 (2005), by permission of the editors. Chapter 5, section 3b, is based on my article "How to Learn a Language like a Chimpanzee", *Philosophical Psychology* 3 (1990), by permission of Taylor and Francis Ltd. Figure 11 in chapter 6 is adapted from Amos Tversky, "Features of Similarity", *Psychological Review* 84 (1977), by permission of the American Psychological Association.

I received very useful comments on an earlier draft from two anonymous referees for Oxford University Press, which led to many improvements in the final version. I also thank Willem A. de Vries, Giovanni Mion, Matthew Van Cleave, and Franklin Scott for helpful comments on various parts of the draft.

This book exists now, not later, due to the largesse of the Taft Research Center of the University of Cincinnati, which, through a grant to my department, enabled me to spend the entire academic year 2008–9 piecing this book together and which, through a grant to me for the summer of 2010, allowed me to add the final touches. I am grateful to the officers of the Taft Research Center, my then department head, John Bickle, and the Dean of the College of Arts and Sciences, Valerie Hardcastle, for those opportunities.

Introduction

Defining the Question

In this book, I propose to explain how ideas such as *dog* and *chair* arise in the mind. So my topic is ordinary ideas, such as *dog* and *chair*, not big ideas such as *democracy*, *science* or *the intermittent windshield wiper*. The traditional view has been that ideas, or, as I will call them, *concepts*, are somehow extracted from perceptual experience. I will argue that this tradition is mistaken. A great deal of problem-solving can be achieved by means of a form of imagistic thinking that does not involve the application of concepts at all. Included in the kind of problem-solving that this nonconceptual mode of cognition makes possible is language learning and word choice. In view of this fact, conceptual thought can be identified with the use of the very languages we speak, and concept formation can be equated with language acquisition.

Concepts may be defined as the building blocks of judgments. When you look at a furry animal, you may judge that it is a dog. When you have finished your gardening, you may judge that all of the daffodil bulbs have been planted. The concepts *dog* and *daffodil* and *planted* are components of these judgments. Thus, conceptual thought involves the classification of things as belonging to kinds, as when we judge of some object that it is a dog or judge that something has been planted. In contrast, perceptual experience, and imagistic thought more generally, draw no functional boundaries between one kind of thing and another kind. Imagistic thinking rests, instead, on an ability to track things through space, on an understanding of how one event can lead to another, and on a nonconceptual capacity to recognize that an object *x* is more like an object *y* than like an object *z*. Concepts, I will argue, cannot be extracted from perceptual experience, and imagistic thought does not depend on classification.

Traditionally, linguistic communication has been treated as a means by which a speaker reveals to a hearer the conceptual content of an underlying thought. In my opinion, this conception of linguistic communication is a mistake. An alternative, which I defend, is to think of linguistic communication as a means by which people coordinate their actions. The cognitive processes by which interlocutors decide what to say are, at the most basic level, processes of imagistic thought. Imagistic thought represents whole

objects and events that transpire among them. Imagistic thoughts are not *what is communicated*, but imagistic thought is the means by which we use language cooperatively. By talking, interlocutors construct a linguistic representation of the *context* pertinent to their conversation. What they do in pursuit of the goals of their conversation depends on what they take the context pertinent to their conversation to be. Once agents have acquired basic linguistic skills, they can productively talk to themselves. In this light, it makes sense to think of spoken language as the medium of conceptual thought.

In saying that *spoken* language is the medium of conceptual thought, I do not mean that conceptual thinking is confined to people who can *speak*. English is a spoken language, but some users of English cannot speak it but can only write it. I also mean to include, as possible media of conceptual thought, languages that are not literally spoken at all, such as varieties of sign language. So when I refer to *spoken* language, the term is to be understood in a broad sense, as including all languages by means of which people communicate with one another. I call it *spoken* language to distinguish it from the so-called language of thought that some theorists have posited in order to explain conceptual thought.

1. What are concepts?

I said that concepts are the building blocks of judgments; but that statement picks up our subject somewhere in the middle. So now I want to address in a more fundamental way the question: What is our subject matter when we talk about the nature of concepts and conceptual thought? In my opinion, the subject matter is best identified by way of a certain traditional (but in my opinion mistaken) theory of linguistic communication.

The traditional view is that linguistic communication is basically a matter of a speaker's speaking words that will enable the hearer to recognize that the speaker has in mind a certain thought (e.g. Davidson 1990; Jackendoff 1995; Searle 2007, and countless others). The thought that the speaker intended the hearer to recognize in the speaker on the basis of the speaker's choice of words is the one that the speaker *expressed*. If the speaker's act of speech is informative and the hearer *accepts* what the speaker says, then the result will be that the hearer comes to have the *same* thought as that which the speaker expressed. The sort of thought expressed when a speaker makes an *assertion* using a sentence in declarative mood will be a *judgment*.

For example, if you say, "Some mammals lay eggs", then on the basis of my understanding of the English language (and the assumption, perhaps only tacit, that you are speaking English), I may infer that you judge *that*

some mammals lay eggs, and that, presumably, is at least one thing you intend me to do. So we can say that you *express* the judgment that some mammals lay eggs. If I did not already believe that some mammals lay eggs, and, for whatever reason, I am prepared to accept the judgment that you express, then the result will be that I too judge that some mammals lay eggs. You may have intended that further result, or, if you did not expect me to take your word for it, you may have intended only that I would recognize that *you judge* that some mammals lay eggs. You may also have intended other results. You may have intended that I recognize that you are thinking of platypuses. Or you may have intended that I conclude that not only birds and reptiles lay eggs, or may have intended that I marvel at the diversity of animal life. But by being careful about the *means* by which a speaker expects to achieve such results, we should be able to distinguish the judgment expressed from other thoughts the speaker may have intended the hearer to recognize in the speaker and to distinguish that recognition of what is expressed from other effects that the speaker may have intended to produce in the hearer.

We expect a certain correspondence between the structure of a sentence and the structure of the judgment that an assertion made by means of that sentence expresses. So if a judgment is expressed by means of the sentence "Some mammals lay eggs", then there will be a component of the judgment expressed that corresponds to the word "mammal", a component that corresponds to the word "eggs", and a component that corresponds to the word "lay". Instead of *component* we could say *aspect*, to allow that a judgment, as it is embodied in the brain, is an indivisible whole having a variety of *properties*, but for simplicity I will always say "component". The word "some" is perhaps a different case. We might think of that word as part of the syntactic structure or, instead, as a *mode* of combining the others, in which case, what corresponds to it in the judgment might be better described as a feature of the *way* in which the others are combined. In general, we can say that for each open class word in a sentence (the nouns, verbs, adjectives, etc., but not the particles or logical words) there is a corresponding component in the judgment that the sentence can be used to express.

In light of this correspondence we may refine our characterization of concepts as the building blocks of judgments. For each judgment there will be a sentence that can be used to make an assertion that expresses it, and that judgment will have components that stand to the whole judgment as the open class words of the sentence that can be used to express it stand to the whole sentence. Concepts are *such* components of judgments. In saying this, we do not *define* concepts in terms of language, but we use a certain conception of language to locate the subject matter of a theory of concepts.

In describing concepts in this way I am assuming that judgments do divide into components (or aspects) that can be recombined to form further judgments, just as the words in a sentence can be recombined to form other sentences. As we expect assertions to be the linguistic expression of judgments, we expect the structure of the judgment expressed to shape the speaker's choice of words in his or her assertion. Since the same words may occur in the expressions of different judgments, we expect that different judgments will have components that are the same when the words that express those components are the same. As a consequence of the fact that we expect the words that express a judgment to be guided by the structure of the judgment expressed, we expect our capacity to form judgments to exhibit a certain amount of what is called *systematicity* (Fodor and Pylyshyn 1988). For example (from Fodor and Pylyshyn), a capacity to judge that John loves the girl entails a capacity to judge that the girl loves John. (But this capacity can be exaggerated. A capacity to judge that John fell in the lake need not entail a capacity to judge that the lake fell in John.)

In defining concepts as the building blocks of judgments I do not deny that concepts also occur in wishes and wonderings and other sorts of mental state. But I suppose that the concepts that go together to form wishes and wonderings also form the building blocks of judgments. So, as the concept *egg* may be a component of my wondering whether there are any eggs in the refrigerator, that same concept may be a component in my judgment that, yes, there are three eggs in the refrigerator.

Sometimes people assume that by "concept" one must mean *abstract concept*, and assume that abstract concepts represent something other than the perceptible properties of things. According to this scheme, a thought to the effect that *this is green*, or a thought to the effect that *this is triangular* might not qualify as conceptual. On the contrary, these judgments are just as much conceptual thoughts as the judgment that some mammals lay eggs. These are judgments to the effect that a particular object has some property. The difference is at most that the property in question might be one that we can recognize in a thing in some sense "directly", without bringing to bear any background knowledge.

Locating our subject matter in this way in terms of a theory of linguistic communication does not entail that concepts are possible only for creatures that speak a language. So in introducing the concept *concept* in this way, I am not already taking a step toward my thesis that language is the very medium of conceptual thought. We can consistently maintain that concepts are the kinds of thing that compose the judgments that we express in words without supposing that every creature that can form such judgments possesses as well the capacity to express them in words. The concepts of such a creature may be in various ways more primitive than the concepts

of language speakers. Still, they might qualify as concepts inasmuch as they are the same basic kind of representation as that which we find in language speakers, which go together to form mental acts recognizable as judgments.

One question this characterization of concepts as the building blocks of judgments does not decide is whether a concept is something in the head of an individual thinker or whether it is something that is in some sense shared between several thinkers. If you judge that some mammals lay eggs and I judge the same, do we have two judgments, the one in you and the one in me? Or is there just one judgment, the one we both make? Judgments can be counted in either way, depending on our purposes, and concepts, considered as the building blocks of judgments, will likewise be countable in these two ways, as I will now explain.

When you utter the sentence "Some mammals lay eggs", there is something in your brain that is the judgment you express. We can think of this thing in your brain as a concrete particular (of a neurological kind). When I accept what you say, there is likewise something in my brain that is my judgment that some mammals lay eggs. The concrete particular that is your judgment is not the concrete particular that is my judgment, because your head is a different location from mine. A concept, considered as a building block of a concrete particular judgment, will likewise be a concrete particular. (Of course, we will discriminate such parts within a particular representation only insofar as we think of these parts as belonging to distinct types of some kind.) So the concept that is the building block of your concrete particular judgment corresponding to the word "mammal" is not the concrete particular in my brain corresponding to the word "mammal". Concepts, considered as components of concrete particular judgments, are not shared.

On the other hand, if you judge that some mammals lay eggs and I too judge that some mammals lay eggs, then inasmuch as both of our judgments are characterizable as judgments *that some mammals lay eggs*, our judgments are in this respect the same. When two judgments, considered as distinct particulars, are the same in a way that allows them to be characterized by means of the same "that"-clause (by a third person referring to the same times, places, and objects), let us say that they have the same *content*. Thus, your judgment and mine have the same content, and in this respect, they are *the same judgment*. It is commonplace to speak of content Platonistically, as if it were an object in Plato's heaven, distinct from things in brains. Speaking in that way, we may say that your judgment and my judgment *possess* or *bear* or *express* the same content. (Expression in this sense is something different from the relation of expression that I spoke of above, which is a relation between a judgment and an act of speech.) But we

do not have to take this Platonistic way of talking literally, and we can understand the sameness of content as a certain kind of sameness between particular judgments in different people's heads. The term "concept" can also be used to refer to a component of such a shared content. As such, a concept may be shared; it is a type to which a token in you and a token in me may both belong. Inasmuch as we share the judgment that some mammals lay eggs, we share the concept *mammal*.

When I said, just now, that if you judge that some mammals lay eggs and I judge that some mammals lay eggs, then our judgments have the same content, I did mean *exactly the same*. But what if you are thinking of platypuses and echidnas and I imagine egg-laying mice? Or what if you define a mammal as an animal that has hair and breathes air and I think of a mammal as a warm-blooded animal with sweat glands? Can it still be said that our concepts of mammals are the same? Yes it can, but this may be hard to see, because the phrase "concept of mammal" might also be used to refer to what I will call a *conception*, or *conceptualization*, of mammals. (The confusion of concepts and conceptions has been discussed by Rey 1983 and Cummins 1996, pp. 88–9.) If we both judge that some mammals lay eggs, then we both have the concept *mammal*—the very same one. But we may nonetheless have different *conceptions* of mammals. Our conceptions of mammals are made up of our beliefs and stereotypes. If you believe that all mammals have hair but are not so sure whether they all have sweat glands, whereas I believe that they all have sweat glands but am not so sure that they all have hair, then we have different *conceptions* of mammals even if it is the same concept *mammal* that enters into your belief as enters into mine. Or if for you a bat is a typical mammal, whereas for me it is hard to think of bats as mammals at all, then your conception of mammal may differ from mine. One needs to presume a sameness of concepts even to characterize our differences as differences in *judgment*. In supposing that the judgment that all mammals have hair is in conflict with the judgment that some mammals do not have hair, one supposes that the concept *mammal* is the same in both judgments.

The reason it makes sense to say that two people have the same concept *x* despite their differing conceptions of *x*'s is that we may identify our subject matter, concepts, in terms of a conception of linguistic communication, as I have proposed. If we want to characterize successful linguistic communication as I did above, then we need a concept *concept* that allows two people to have the same concepts despite big differences in their conceptions. Despite your thinking of platypuses and my thinking of mice, I may understand you perfectly well when you say to me "Some mammals lay eggs". According to the theory, that understanding consists in my recognizing that you are expressing the judgment that some mammals

lay eggs. I can *recognize* that you have a judgment that you actually have only if I possess the concepts that make up the judgment that I recognize in you. The fact that two people with different conceptions of *x*'s can possess the same concept *x* does, I acknowledge, create difficulties when we go to try to say what concepts *are* (see chapter 4).

So far I have defined *expression* in terms of communication; I have defined *judgments* as what assertions express; I have defined the *content* of judgments in terms of "that"-clauses; and I have defined *concepts*, considered as types, in terms of the contents shared in communication. These definitions fit together only loosely, inasmuch as the content we think of as communicated is not always perfectly reflected in the words that make up the "that"-clause by which we describe the judgment expressed. For example, if someone asserts, "Chick peas are garbanzo beans", and thereby expresses the judgment *that* chick peas *are* garbanzo beans, we might still be uncertain whether the concept that corresponds to "chick peas" in the content of that person's judgment is the same as that which corresponds to "garbanzo beans". Or if I think that Obama *is tall* and Milly thinks that her third-grade classmate *is tall*, is it the same concept of tallness that we apply in these two cases? In one sense, yes, and in another, no. However, for present purposes it will not be necessary to produce definitions that mesh more precisely.¹

The literature on the nature of concepts has not always treated them as the building blocks of judgments. There are theories of concepts that positively exclude the possibility that concepts are the building blocks of judgments. For instance, if we analyze concepts as *abilities* (cf. Millikan 2000), then it will be just a category error to say that concepts are the building blocks of judgments, since an ability cannot be a building block of an event or a state such as judgment. Or if we analyze concepts as knowledge structures (cf. Keil 1989), then too we will not be able to treat concepts as the building blocks of judgments. On the contrary, just the

¹ In the contemporary literature (e.g. Chalmers 2006), so-called "Russellian contents" are sometimes modeled as structures built up from actual objects and properties. So the Russellian content of "Socrates is snub-nosed" would contain Socrates himself and the property of being snub-nosed as components. If we say that and at the same time say that concepts are the building blocks of contents, then we seem to have to allow that concepts may be identical with individuals and properties. That would be a strange, though not intolerable, conclusion; but I am not adopting it. This same literature identifies so-called "Fregean content" with something so subjective and idiosyncratic that it cannot possibly be identified with the content communicated in linguistic communication. But that is clearly not what Frege himself intended (1994 [1892]), and Fregean content in this non-Fregean sense is not built up from concepts in the sense I have attempted to define.

opposite will be closer to the truth: judgments will be the building blocks of concepts.

However, anyone who accepts a certain commonplace view of the role of concepts in language learning will treat concepts as the building blocks of judgments. The commonplace view is that children learn certain words (such as nouns, verbs, and adjectives) by mapping words into concepts that they form more or less independently. Thus, Paul Bloom writes,

Learning a word involves *mapping* a form, such as the sound “dog,” onto a meaning or concept, such as the concept of dogs. (2002, p. 89)

And according to Gregory Murphy,

... a word gets its *significance* by being connected to a concept or a coherent structure in our conceptual representation of the world. To put it another way, the meaning is built out of concepts. (2002, pp. 388–9)

Or again, according to Eve Clark,

One issue for language acquisition is how children find out which meanings there are words for; another is just how they map each meaning to the right word. (2003, p. 9)

(Clark equates meanings with “conceptual categories”.) The assumption that word learning is a matter of mapping words into concepts is almost universally accepted among psychologists who study language acquisition. The only exceptions might be a few surviving behaviorists and the practitioners of what is called “cognitive linguistics” (such as Leonard Talmy). If we accept this basic conception of language learning, then inevitably we will conceive of concepts as the building blocks of judgments. Words, we will say, can be learned by being mapped into concepts just because judgments are built up from concepts in much the way that the sentences that express them are built up from words, and sentences express judgments.

In this way, I introduce our subject matter, the concept *concept*, via a theory of linguistic communication (which, again, does *not* mean that only language-speaking creatures can have concepts). The theory of linguistic communication that I have appealed to is not vacuous. It makes certain assumptions that may well be false, and which I think are in fact false. In particular, it assumes that judgments and their building blocks, concepts, have the kind of independence from language that they need to have in order that we can appeal to them in explaining linguistic communication. The theory says that a speaker chooses to speak certain words because the speaker *intends* that the hearer will *recognize* that the speaker makes a certain *judgment*. That assumes that things like intending and recognizing

and judging are mental states whose nature we can understand apart from language—that they are not themselves linguistic acts. Probably everyone will agree that the words we hear spoken around us play a role in determining which concepts, of all those that we *might* form, are the ones that we *do* form. But one can allow that much while maintaining that the kind of thing a concept is does not essentially depend on language, and that is what a proponent of the theory of linguistic communication that I have here described must hold.

In the first half of this book, I will criticize all major conceptions of concepts that lend them the necessary independence from language. In thus denying that concepts have the requisite independence from language, I will reject the theory of linguistic communication by which I introduced the concept *concept*. That might look like cutting off the branch I'm sitting on. In general, if one considers a concept that we grasp only in light of a certain theory (here the concept is *concept*), and one then discovers that that theory is mistaken, one has two choices. One can deny that the concept in question applies to anything in reality. That is, one can become an eliminativist about that kind of thing. Or one can try to assimilate the old concept into the new theory, proposing a new understanding of that which one formerly understood in terms of the now rejected theory. The problem with the former approach, in the present instance, is that there does not seem to be any real prospect of teaching people never to use the word “concept” in describing the processes of thought. I am using it myself, and not merely talking about it, in this very paragraph! So I think that I need to assimilate the concept *concept* into the new theory. I propose to do so by identifying conceptual thought with the use of language.

My treating concepts as the building blocks of judgments will remind some contemporary philosophers of the views of Gottlob Frege (1994 [1892]). Frege's term for the sense of a sentence was *thought*. Thoughts, in Frege's sense, were not only the contents of judgments; they could be the contents of wonderings and other attitudes just as well. He thought of thoughts as what people communicate by means of language, and he thought of them as the *cognitive values* of sentences. Thoughts in Frege's sense were compounded of other objects that were the senses of subsentential expressions—the sense of a name, the sense of a verb. These senses of words and phrases are roughly what I am calling concepts, considered as shareable. (Confusingly, what Frege himself called “concepts”—*Begriffe* in German—belonged not to the realm of Fregean senses but rather to the realm of referents.) Frege was not the first to think of thoughts in this way, as distinct from sentences but composed of elements in the manner of sentences; but he may be the source uppermost in the minds of many contemporary philosophers. Frege wished to steer us away from

the Kantian tradition, according to which concepts are bound up with imagistic representations. I certainly will not be trying to steer us back that way, and will devote a chapter to criticizing the Kantian tradition myself, even though I will acknowledge a role for imagistic representation in language learning and language production. But, as I have just explained, in the end I will not have a place for anything quite like Frege's conception of thoughts and other senses at all; there is nothing like that that hovers alongside spoken words to serve as their meaning.

2. Philosophers versus psychologists

The enterprise of explicating the nature and origin of concepts belongs as much to psychology as it does to philosophy. However, there are some significant differences between the questions that psychologists ask and the questions that philosophers ask. What the psychologists who study concepts are primarily interested in doing is describing a mechanism in terms of which we can explain what people do in certain experimental situations (in hopes that this will explain as well what people do in the wild). What the philosophers want to know is what the relation is between the building blocks of judgments, on the one side, and things and properties in the world, on the other, such that in terms of that relation (the reference relation) we can explain the conditions under which a judgment is *true*. These different interests are so different and lead to such different theories that one might question whether the psychologists and the philosophers are really talking about the same thing.

A typical question for a psychologist would be: Why do people more quickly label a robin as a "bird" than they label a penguin as a "bird"? A typical answer will be that the word "bird" is associated with the concept *bird* and the concept has some kind of internal structure that allows it to be applied more quickly to a robin than to a penguin. For example, the concept might include descriptions of a typical bird's features, which a robin possesses and a penguin lacks. Or a psychologist might observe that children understand that a wolf in sheep's clothing is still a wolf and conclude that children's applications of concepts are not based exclusively on external appearance but also utilize children's general theories.

A typical question for a philosopher would be: What is the relation of reference that holds between a certain mental representation *a* and a certain particular object in the world and between a certain representation *F* and a certain property of things in the world such that the representation *F(a)* is *true* if and only if the object has the property? The representation *a* might be a kind of mental demonstrative, an inner "that", and the representation *F* might be a concept, such as the concept *bird*, in which case the

representation $F(a)$ is true if and only if the object mentally demonstrated has the property of being a bird. (I should stress that a reference relation is a relation between something in the mind and something *outside* of it. It is emphatically *not* a relation of association between a symbolic representation and some other kind of representation.)

The psychologist's question and the philosopher's question are very different. An answer to the philosopher's question will not answer the psychologist's question. A penguin is no less a bird than a robin. So if F refers to the property of being a bird, then $F(a)$ is true in the same way when a refers to a particular penguin and when a refers to a particular robin. And an answer to the psychologist's question will not answer the philosopher's question. If F refers to the property of being a reptile, then if a refers to a particular chameleon, $F(a)$ is true; whereas if a refers to a particular pangolin (which is a mammal), then $F(a)$ is false. But a pangolin looks a lot like a reptile. So the internal structure of my concept *reptile*, such as the psychologist describes, may render it as applicable to the pangolin as it is to the chameleon when I have only perception to go by. So the psychologist's theory will not explain the difference between truth and falsehood that interests the philosopher.

And yet, as I will show as we proceed through the various types of theory, the philosopher and the psychologist *are* talking about the same topic. My introduction of the topic in terms of a theory of linguistic communication helps us to see that they are. As far as I know, every psychologist working in the field of concepts intends his or her theory to be a central player in an account of language learning and word choice. Often, moreover, psychologists offer specific accounts of how sentence-building devices (such as the quantifier "all" or the sentential connective "or") map into the sorts of structures they have described. It is true that the theoretical entities posited by psychologists do not often include anything quite like a *judgment expressed* as I have described it. But what they say about language is generally enough to establish that what they are talking about in explaining categorization behavior is supposed to be the very things the philosophers are talking about in explaining truth.

So I think that Edouard Machery is mistaken when, in a recent book, he concludes that philosophers' and psychologists' theories of concepts simply deal with different topics (2009, p. 32; see also Piccinini and Scott 2006). Machery entertains the thesis that for both parties concepts are the building blocks of judgment and rejects it on the grounds that what it means is "not fully explained" (2009, p. 26). But that is not a reason to reject it, because what we want for purposes of characterizing a wide range of theories is precisely a conception of concepts that we can recognize in different manifestations apart from theoretical details. We all have an intuitive

sense of what it is to make an assertion, and insofar as our assertions express judgments, we can identify the building blocks as parts that stand to the whole judgment as words stand to sentences.

According to Machery, the conception of concepts that psychologists share is that of a “body of knowledge” stored in long-term memory that is “used by default” in higher cognitive processes (2009, p. 12). But that is surely not a concept of concepts that we have any grasp on apart from some particular psychological theory, and I do not find that psychological thinking about concepts does reliably have a use for Machery’s notion of a *default*. Psychological studies described as studies of concepts, or “categories”, have sometimes been studies of what Machery calls “bodies of knowledge” (a good example would be Keil 1989; for substantiation, see my review, Gauker 1991b). That does not mean that the object of their study was not concepts as I have defined them; rather, these researchers have sometimes confused concepts and conceptions. A good deal of what passes for the study of concepts would be better conceived as a study of the representational structures that may be used in deciding whether a concept applies in a given case.²

One could grant that there is a core conception of concepts that virtually all researchers share but still maintain that different kinds of psychological explanation posit various kinds of mental representations and that some of these deserve to be called *concepts* though they are not the building blocks of judgments. I will not argue directly against that, although at several points I will offer alternative explanations of phenomena that some might have wished to explain in terms of such “concepts” (chapters 3 and 5). In my opinion, the characterization of the phenomena as requiring the application of *concepts* is often not warranted because it is not shown that the representations involved utilize functional boundaries between the kinds purportedly represented. If other kinds of mental entities deserve to be called concepts that are not the building blocks of judgments, then my thesis is simply not concerned with them. However, I urge readers to guard against too quickly dismissing my critical arguments as not touching the topics that interest them.

² Weiskopf’s (2009) critique of Machery’s view is revealing. While Weiskopf holds that what psychologists call concepts are many kinds of thing (“prototypes, bundles of exemplars, theory-like structures of some sort”, p. 145), he also holds that all of this variety deserves to be called concepts because they have certain things in common. At the top of his list is the fact that concepts enter into mental states having logical structure and entering into logical inferences. Thus Weiskopf in effect agrees with me that concepts are essentially the building blocks of judgments.

3. My theses

Some philosophers have argued that language is necessary for certain kinds of thought but have not gone so far as to claim that language is the very medium of all conceptual thought. Bermúdez (2003a) has claimed that language is necessary for a certain sort of reflective thought, which in turn is necessary for full-fledged logical thinking. Camp (2009) allows that language may be necessary for a certain highly general capacity to generate thoughts through recombination of conceptual components. However, Bermúdez and Camp also countenance intermediate varieties of thought that they think of as bearing conceptual content. I go much further: Language is necessary for every kind of thought that involves judging, of some particular, that it belongs to some kind.

I do not mean by this that conceptual thought without spoken language is inconceivable or metaphysically impossible. Here is one way to conceive of it: Imagine a community of people who speak a language that does not in fact exist. Next, imagine a person who learns this language in the normal way, but then at some point in his life decides never to speak out loud again, although he continues sometimes to think in that language. Now imagine that a duplicate of such a person, full grown, in the state he is in after his vow of silence, coalesces by cosmic accident somewhere in a hospitable environment in a world where that language has never been spoken. There I suppose you have conceived of a creature capable of conceptual thought who does not speak a spoken language, or at least a perfect simulacrum of one. My claim is only that there is no creature on earth that thinks conceptually but does not use a spoken language to do it, nor is there likely ever to be one.

The argument for this strong language-dependence thesis will be, first, that no language-independent account of how concepts arise in the mind has ever succeeded and, second, that we can hope to explain how language arises in the mind on the basis of a kind of nonconceptual, imagistic thought. That no other theory has ever succeeded will be the burden of the first four chapters. Concepts do not arise through abstraction from perceptual experience or as principles for organizing perceptual experience, and yet they are not all definable in terms of a repertoire of innate conceptual primitives. Of course, I will be able to examine closely only a representative sample of attempts. I do not have any *a priori* argument to show that no language-independent account of concepts *could* succeed. Insofar as I despair of all language-independent accounts of the origin of concepts, my inference to this despair is inductive. By categorizing theories according to their historical roots and then attacking each type at the root,

I hope to persuade the reader that no viable theories are likely to sprout from this same soil. My hope is that these chapters will persuade the reader that we desperately need to make a radical break with the paradigms that have dominated our thinking about concept formation in the past.

In the second half of the book I will first describe a kind of nonconceptual, imagistic cognition and argue that it does not rest on conceptual thought, and then I will argue that a simple language can be learned by means of this kind of imagistic cognition. Mental imagery, as I define it, represents not sensory qualities but whole objects and scenarios and motions and changes among them. Having characterized a kind of imagistic cognition, I certainly will not go on to claim that concepts arise from it in just the ways I will have denied in the previous chapters. So I will not claim images are meanings or that meanings are somehow grounded in images. Rather, my claim will be that spoken language can be added to imagistic thinking in a way that facilitates interpersonal cooperation. Finally, I will put forward my proposal that we may identify conceptual thought with the use of language. At that point I will take up the question of what it could possibly mean to “think in language”.

4. Some disclaimers

In view of the positive account of conceptual thinking that I will offer, this book is as much a thesis about the nature of language as it is a thesis about the nature of concepts. That positive thesis rests on a negative thesis against the traditional conception of linguistic communication that I have described above. This book constitutes part of the argument against that traditional conception inasmuch as it criticizes a conception of the origin of concepts on which the traditional conception relies. There are other parts of the argument that are not developed here. Here I do not take up the theories of contentfulness and the theories of interpretation on which the traditional conception relies. (For that, see my book *Words without Meaning*, 2003a.)

Books on the nature of concepts typically spend a lot of words on the differences between the classical, the prototype, and exemplar theories, and may add a discussion of the “theory”-theory of concepts (Smith and Medin 1981; Murphy 2002; Machery 2009). Many of these theories fall within the spectrum of views surveyed in the first four chapters of this book, but I will focus on the work of particular, representative authors rather than on these categories. Still, this book is not intended to be a review of the concepts literature, and it is not serviceable as one. Some of the topics commonly

addressed in such reviews, such as conceptual combination and ad hoc concepts, do not come up at all.

My aim is to demonstrate a possibility. I want to show that there is a kind of imagistic thinking that does not rest on concepts, that it can do a lot, and that in terms of it we might be able to explain how a language might be learned. Showing that language can be learned on the basis of nonconceptual thought will free us up to conceive of spoken language as the very medium of conceptual thought. What is not necessary for that project will not be part of this book. Consequently, I will not have anything to say about the child's course of conceptual development. Not only do children learn many facts as they grow older, they evidently also become better thinkers and problem-solvers. In view of the fact that children of a given age will make the same sorts of mistakes as other children of that age, there appears to be a pattern of development that one would like to capture in a psychological theory (Carey 2009, ch. 10). However, doing that is not part of the plan for this book.

Even though the subtitle is "An Essay on the Origin of Ideas", I will not have anything to say about how conceptual thought arises for the first time in a society or in a species. My topic is only how concepts arise in minds embedded in societies where other people already have them. If we can explain how one mind picks up on concepts that are already present in the minds of those around them, then it should be possible to extend that explanation to an account of how concepts arise in the first place. The story will be much the same except that accidents will be responsible for much of what happens in a more regular and predictable way when a member of a society picks up on an idea from others who already have it. Granted, this is a little like saying that once we understand biological reproduction, we will have most of what we need in order to understand biological evolution.

Finally, I want to disavow all taint of *Whorfianism*. Whorfianism (named after the anthropologist Benjamin Lee Whorf) is the thesis that people's thinking is shaped by the nature of the languages they speak. For instance, it might be claimed that people who speak a *counter language*, such as Korean or Tzeltal, think of the world as consisting of stuffs and think of individuals, even individual people, as merely discrete lumps of that stuff (cf. Imai and Gentner 1997 for a more cautious hypothesis in this direction). Anyone who claims that spoken language is the very medium of conceptual thought, as I do, is liable to be accused of Whorfianism. Thus, Lila Gleitman and Anna Papafragou write: "Do our thoughts take place in natural language? If so, it would immediately follow that Whorf was right all along" (Gleitman and Papafragou 2005, p. 636). Gleitman and Papafragou are simply wrong about this. I am prepared to allow that differences between

languages make small differences in the way their speakers think. But Whorfianism is certainly not entailed by my thesis, and I am not a Whorfian, because one can maintain that language is the medium of conceptual thought, as I do, while also maintaining that the world's languages are all enough alike that people who think by means of any of them think basically alike. (I return to this topic in chapter 8.)

Chapter 1

The Lockean Theory

Most of our concepts, it is commonly supposed, rise up out of our sensory perceptions. By means of our senses we perceive the objects around us. The product of perceiving is perceptions. Perceptions are certain sorts of objects in the mind; by means of them we perceive things in the world around us. The mind performs various operations on these mental objects, and the product of these operations is concepts.

What are these perceptions? What do they represent? Do they represent particular things such as *this table* and *that dog*? Or do they represent only certain properties that objects have, such as their colors and shapes? And what are these mental operations performed on them that generate concepts? Are they some kind of putting together or some kind of taking apart?

Some early answers to these questions may be found in John Locke's *Essay Concerning Human Understanding* (first published in 1689). (Locke lived from 1632 to 1704.) His views continue to be influential, although his influence is often obscured through a mischaracterization of what he actually said. Locke's text actually contains *three different* theories of concept formation that he did not clearly distinguish but which we can distinguish for him.¹ As I will illustrate, citing the work of psychologists Eleanor Rosch and Jean Mandler and philosopher Jesse Prinz, each of his theories still occupies a place in the contemporary psychological and philosophical literature.

Where I speak of concepts, Locke speaks of *ideas*. I will assume that Locke's theory of ideas can be taken as a theory of concepts in my sense. Evidence for this is that Locke holds that words, such as nouns and verbs, "signify" ideas in the mind of the speaker (III, ii, 1–2). In using the word "signify" in this way, he certainly does not mean that what we *talk* about is only our own ideas; probably he means to emphasize that the first thing one has to attend to in making sure that one speaks intelligibly is the ideas one has in mind (III, xi, 8). More important, for my purposes, is Locke's account of the function of speech:

¹ Citations to Locke's text will follow standard practice in listing first the number of the book, in roman numerals, second, the number of the chapter, in lower case roman numerals, and third, the section numbers, in Arabic numerals.

When a Man speaks to another, it is that he may be understood; and the end of Speech is, that those Sounds, or Marks, may make known his *Ideas* to the Hearer. (III, ii, 2)

Locke was not in a good position to represent the hearer as drawing an inference from the speaker's choice of words to the content of an underlying thought. Locke was not very conscious of the possibility of black box reasoning; otherwise he might not have been so pessimistic about the possibility of discovering the fine structure of matter (IV, iii). Nor does he anywhere have anything to say about *ideas of ideas*, such as he would need to countenance if he were to have an account of how one mind contemplates the ideas in the mind of another. So he was in a poor position to think of the hearer as drawing an inference from the speaker's choice of words to the content of an underlying thought. He was in no position to explicate *expressing* an idea as a matter of the speaker's intending the hearer to infer that the idea is present in the mind of the speaker, as contemporary philosophers of language might do.² Instead, for Locke the way in which words reveal the ideas in the speaker that they stand for is by *exciting* in the hearer those same ideas.

[Words] being immediately the Signs of Mens *Ideas*; and, by that means, the Instruments whereby Men communicate their Conceptions, and express to one another those Thoughts and Imaginations, they have within their own Breasts, *there comes by constant use*, to be such a *Connexion between certain Sounds, and the Ideas they stand for*, that the Names heard, almost as readily excite certain *Ideas*, as if the Objects themselves, which are apt to produce them, did actually affect the Senses. (III, ii, 6)

So it is fair to describe Locke as holding that words *express* ideas, although he did not explicate the relation of expression in the way a contemporary philosopher might do.

As I say, Locke seems to have at least three different theories of how ideas arise in the mind. In all of these, he is concerned to deny that ideas are innate; in all of them, ideas arise somehow out of sensation or out of reflection on one's own mind. One of Locke's theories is what I will call the *composition theory*. According to this, what enter the mind initially are ideas of certain perceptible qualities, such as color and shape, and these are *combined* to form ideas of kinds of object, such as *gold*. Another is what

² For these reasons I am not quite in agreement with Ott (2004), who holds that for Locke words are *indicators* of ideas. That answer leaves us wondering how Locke could countenance an indication relation between two kinds of thing, one of which (ideas in the mind of another) was imperceivable. Locke does, at one point (II, xxix, 12) speak of "divining" the idea that a name stands for in the mind of another man, but he has nothing to say about how this is done.

I will call the *abstraction-as-subtraction theory*. According to this, what enter the mind initially are ideas of particular objects, such as the idea of Peter and the idea of Mary; by a process of subtracting what differentiates these from one another, the mind *abstracts* ideas of the kinds that these particulars belong to, such as the idea *person*. Finally, there is a third theory, much less prominent in the text, that I will call the *abstraction-as-representation theory*. According to this, the mind contains only ideas of particular qualities or particular objects, but some of these may be treated as representing in our thinking a wide range of other things.

Offhand, Locke's three theories are simply incompatible with one another; they do not merely highlight different aspects of a single conception of ideas. In the main text of this chapter, that is how I will treat them. However, if our interpretation of an important author finds him simply inconsistent, that is a reason to doubt the interpretation. So in the appendix to this chapter I will try to bolster my attribution of these three theories to Locke by showing that if we are willing to attribute to Locke some assumptions that he himself does not explicitly state, then there may be a way to unify the three theories.

In this chapter, I will set out each of Locke's several theories in detail and criticize each one. Then I will examine contemporary versions of Locke's several theories and will argue that they are no more successful. Locke's theory of ideas is a good representative of what is taken for common sense about the nature of the mind—perhaps only because Locke's philosophy has been so influential. So it will be edifying to discover that it stumbles on many inconsistencies and that much of contemporary psychology makes hardly any advance beyond Locke.

1. The composition theory

The dominant theme in Locke's theory of ideas is the proposition that simple ideas enter the mind via the senses and the mind forms further ideas by composing these. Through composition the mind forms the ideas of substances, such as gold, and ideas of kinds of thing, such as swans. What I am calling the composition theory is the claim that general ideas are formed by the composition of simple ideas. It is not the theory that Locke sets out when he explicitly undertakes to explain how general ideas are formed—that is the abstraction-as-subtraction theory; but it is clear that he does often treat what are in fact general ideas as the product of such composition.

Simple ideas, Locke tells us, enter the mind “unmixed” (II, ii, 1). The mind is entirely passive in receiving them (II, i, 25; II, xii, 1). If one touches a piece of ice, one passively receives the simple ideas of coldness and hardness. But not all simple ideas are ideas of determinate sensory qualities.