



# LATENT VARIABLE MODELS

An Introduction to Factor, Path,  
and Structural Equation Analysis

FIFTH EDITION

JOHN C. LOEHLIN and  
A. ALEXANDER BEAUJEAN

# Latent Variable Models

An Introduction to Factor, Path,  
and Structural Equation Analysis

*Fifth Edition*

*Latent Variable Models: An Introduction to Factor, Path, and Structural Equation Analysis* introduces latent variable models by utilizing path diagrams to explain the relationships in the models. This approach helps less mathematically-inclined readers to grasp the underlying relations among path analysis, factor analysis, and structural equation modeling, and to set up and carry out such analyses.

This revised and expanded fifth edition again contains key chapters on path analysis, structural equation models, and exploratory factor analysis. In addition, it contains new material on composite reliability, models with categorical data, the minimum average partial procedure, bi-factor models, and communicating about latent variable models. The informal writing style and the numerous illustrative examples make the book accessible to readers of varying backgrounds. Notes at the end of each chapter expand the discussion and provide additional technical detail and references. Moreover, most chapters contain an extended example in which the authors work through one of the chapter's examples in detail to aid readers in conducting similar analyses with their own data. The book and accompanying website provide all of the data for the book's examples as well as syntax from latent variable programs so readers can replicate the analyses. The book can be used with any of a variety of computer programs, but special attention is paid to LISREL and R.

An important resource for advanced students and researchers in numerous disciplines in the behavioral sciences, education, business, and health sciences, *Latent Variable Models* is a practical and readable reference for those seeking to understand or conduct an analysis using latent variables.

**John C. Loehlin** is Professor Emeritus of Psychology and Computer Science at the University of Texas at Austin. He received his PhD in Psychology from the University of California (Berkeley).

**A. Alexander Beaujean** is an Associate Professor of Educational Psychology at Baylor University. He received PhDs in Educational Psychology and School Psychology from the University of Missouri.

This page intentionally left blank

# **Latent Variable Models**

An Introduction to Factor, Path,  
and Structural Equation Analysis

*Fifth Edition*

**John C. Loehlin**  
**A. Alexander Beaujean**

 **Routledge**  
Taylor & Francis Group  
NEW YORK AND LONDON

First published 2017  
by Routledge  
711 Third Avenue, New York, NY 10017

and by Routledge  
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

*Routledge is an imprint of the Taylor & Francis Group, an informa business*

© 2017 Taylor & Francis

The right of John C. Loehlin and A. Alexander Beaujean to be identified as author of this work has been asserted by them in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

*Publisher's Note*

This book has been prepared from camera-ready copy provided by the authors.

*Trademark notice:* Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

*Library of Congress Cataloging in Publication Data*

Names: Loehlin, John C. | Beaujean, A. Alexander.

Title: Latent variable models: an introduction to factor, path, and structural equation analysis.

Description: Fifth edition / John C. Loehlin, A. Alexander Beaujean. | New York, NY : Routledge, 2017. |

Includes bibliographical references and index.

Identifiers: LCCN 2016038052 | ISBN 9781138916067 (hb : alk. paper) |

ISBN 9781138916074 (pb : alk. paper) | ISBN 9781315643199 (epub)

Subjects: LCSH: Latent variables. | Latent structure analysis. | Factor analysis. |

Path analysis (Statistics) | Structural equation modeling.

Classification: LCC QA278.6 .L64 2017 | DDC 519.5/35--dc23

LC record available at <https://lcn.loc.gov/2016038052>

ISBN: 978-1-138-91606-7 (hbk)

ISBN: 978-1-138-91607-4 (pbk)

ISBN: 978-1-315-64319-9 (ebk)

Typeset in Helvetica  
by John C. Loehlin and A. Alexander Beaujean

Visit the companion website: [routledge.com/cw/loehlin](http://routledge.com/cw/loehlin)

# Contents

<b>Preface</b>	<b>ix</b>
<b>Chapter 1: Path Models in Factor, Path, and Structural Equation Analysis</b>	<b>1</b>
Path Diagrams	2
Path Analysis	7
Factor Models	16
Structural Equations	23
Original and Standardized Variables	24
Manifest Versus Latent Variable Models	28
Extended Example	28
Notes	31
Exercises	33
<b>Chapter 2: Fitting Path Models</b>	<b>37</b>
Iterative Solution of Path Equations	37
Matrix Formulation of Path Models	42
Full-Fledged Model-Fitting Programs	46
Fit Functions	54
Hierarchical $\chi^2$ Tests	63
Descriptive Criteria of Model Fit	69
The Power to Reject an Incorrect Model	72
Identification	76
Missing Data	78
Correlations Versus Covariances in Model Fitting	82
Extended Example	84
Notes	86
Exercises	91
<b>Chapter 3: Fitting Path and Structural Models to Data from a Single Group on a Single Occasion</b>	<b>95</b>
Structural and Measurement Models	95
Confirmatory Factor Analysis	100
Some Psychometric Applications of Path and Structural Models	103
Structural Models—Controlling Extraneous Variables	111
Models with Reciprocal Influences and Correlated Errors	115
Nonlinear Effects Among Latent Variables	120
Extended Example	124

Notes	126
Exercises	129
<b>Chapter 4: Fitting Models Involving Repeated Measures, Multiple Groups, or Means</b>	<b>133</b>
Models of Events Over Time	133
Models Comparing Different Groups	142
Fitting Models to Means as well as Covariances	151
The Versatility of Multiple-Group Designs	159
Models with Categorical Indicators	160
A Concluding Comment	163
Extended Example	163
Notes	166
Exercises	169
<b>Chapter 5: Exploratory Factor Analysis—Basics</b>	<b>171</b>
Factor Extraction	174
Estimating Communalities	179
Determining the Number of Factors	183
Rotation	189
An Example: Thurstone's Box Problem	197
Factor Analysis Using Packaged Programs	201
Extended Example	204
Notes	205
Exercises	208
<b>Chapter 6: Exploratory Factor Analysis—Elaborations</b>	<b>211</b>
Rescalings—Alpha and Canonical Factors	211
Alternative Stopping Criteria	214
Alternative Rotation Methods	217
Estimating Factor Scores	220
Hierarchical Factors	225
Nonlinear Factor Analysis	232
Extended Example	235
Notes	240
Exercises	241
<b>Chapter 7: Issues in the Application of Latent Variable Models</b>	<b>243</b>
Exploratory Modification of a Model	243
Alternative Models	248
Can Path Diagrams be Constructed Automatically?	252
Modes of Latent Variable Analysis	255
Advanced Topics in Latent Variable Models	260
Criticisms of Latent Variable Modeling	263

Notes	268
Exercises	271
<b>Appendices</b>	<b>273</b>
A. Simple Matrix Operations	273
B. Derivation of Matrix Version of Path Equations	281
C. LISREL Matrices and Examples	284
D. Model Fit Indices	289
E. Table of Chi-square Values	299
F. Noncentral Chi-square Values for Estimating Power	301
G. Power of a Test of Poor Fit and Sample Sizes Needed for Power	303
H. Communicating About Latent Variable Models	305
<b>Answers to Exercises</b>	<b>309</b>
<b>References</b>	<b>321</b>
<b>Index</b>	<b>367</b>

This page intentionally left blank

## Preface to the First Edition

This book is intended as an introduction to an exciting growth area in social science methodology—the use of multiple-latent-variable models. Psychologists and other social scientists have long been familiar with one subvariety of such modeling, factor analysis—more properly, exploratory factor analysis. In recent decades, confirmatory factor analysis, path analysis, and structural equation modeling have come out of specialized niches and are making their bid to become basic tools in the research repertoire of the social scientist, particularly the one who is forced to deal with complex real-life phenomena in the round: the sociologist, the political scientist, the social, educational, clinical, industrial, personality or developmental psychologist, the marketing researcher, and the like.

All these methods are at heart one, as I have tried to emphasize in the chapters to follow. I have used earlier versions of this book in teaching graduate students from psychology and related disciplines, and have found the particular approach used—via path diagrams—to be effective in helping not-too-mathematical students grasp underlying relationships, as opposed to merely going through the motions of running computer programs. In some sections of the book a certain amount of elementary matrix algebra is employed; an appendix on the topic is provided for those who may need help here.

In the interests of accessibility, I have tried to maintain a relatively informal style, and to keep the main text fairly uncluttered with references. The notes at the end of each chapter are intended to provide the serious student with a path into the technical literature, as well as to draw his or her attention to some issues beyond the scope of the basic treatment.

The book is not closely tied to a particular computer program or package, although there is some special attention paid to LISREL. I assume that most users will have access to a latent-variable model-fitting program on the order of LISREL, COSAN, EQS, or MILS, and an exploratory factor analysis package such as those in SPSS, BMDP, or SAS. In some places, a matrix manipulation facility such as that in MINITAB or SAS would be helpful. I have provided some introductory material but have not tried to tell students all they need to know to run actual programs—such information is often local, ephemeral, or both. The instructor should expect to provide some handouts and perhaps a bit of hands-on assistance in getting students started. The reader going it on his or her own will require access to current manuals for the computer programs to be used.

Finally, it gives me great pleasure to acknowledge the help and encouragement that others have provided. Perhaps first credit should go to the students who endured early versions of the manuscript and cheerfully pointed out various errors and obscurities. These brave pioneers included Mike Bailey, Cheryl Beauvais, Alan Bergman, Beth Geer, Steve Gregorich, Priscilla Griffith, Jean Hart, Pam Henderson, Wes Hoover, Vivian Jenkins, Tock Lim, Scott Liu, Jacqueline Lovette, Frank Mulhern, Steve Predmore, Naftali Raz, and Lori Roggman. Among other colleagues who have been kind enough to read and comment on various parts of the manuscript are Carole Holahan, Phil Gough, Maria Pennock-Roman, Peter Bentler, and several anonymous reviewers. I am especially grateful to Jack McArdle for extensive comments on the manuscript as a whole, and to Jack Cohen for his persuasive voice with the publishers. Of course, these persons should not be blamed for any defects that may remain. For one thing, I didn't always take everybody's advice.

I am grateful to the University of Chicago Press, to Multivariate Behavioral Research, and to the Hafner Publishing Co. for permission to reprint or adapt published materials, and to the many previous researchers and writers cited in the book—or, for that matter, not cited—whose contributions have defined this rapidly developing and exciting field.

Finally, I owe a special debt to the members of my family: Jennifer and James, who worked their term papers in around my sessions at the Macintosh, and Marj, who provided unfailing support throughout.

J. C. L.

**Note to the second edition:** Much of the first edition is still here, but a certain amount of new material has been added, some exercises changed, and one topic (multidimensional scaling) dropped to make room. Also, I've tried to make the book more helpful to those who are using programs other than LISREL. I still appreciate the contributions of the people I thanked before. In addition, I am grateful to Peter Bentler, Robert Cudeck, and Jeff Tanaka for their helpful comments on draft material for the present edition, and to the American Mathematical Society for permission to adapt the table in Appendix F.

**Note to the third edition:** It is still the case that more remains than has been changed. What's gone: IPSOL, BMDP, EzPATH, and a few other items supplanted by the march of events in our field. What's new: more SEM programs, more fit indices, many new references, connections to the Internet, more on means, more on power, and, maybe as important as anything, emphasis on the RMSEA and its use in rejecting null hypotheses of poor fit.

I remain grateful to all those I thanked in the first and second editions, and have a good many names to add—people who gave me advice or encouragement, sent me reprints or preprints or programs, spotted errors, answered queries. These helpful persons include: Jim Arbuckle, Kenneth Bollen, Michael Browne, David Burns, Hsin-Yi Chen, Mike Coovert, Stan Gaines, Steve Gregorich, Greg Hancock, David Kaplan, Timothy Keith, Robert MacCallum, Herbert Marsh, Tor Neilands, Frank Norman, Eddie

Oczkowski, Ed Rigdon, Doris Rubio, Bill Shipley, Jim Steiger, Bob Thorndike, and Niels Waller. And if I've left anybody out—well, them, too.

**Note to the fourth edition:** The basic approach of the fourth edition remains the same as that of previous editions, and, mostly, so do the contents of the book, with some mild reorganization. Chapters 3 and 4 are now divided slightly differently, so that Chapter 3 covers single-group, single-occasion models, and Chapter 4 deals just with models involving multiple groups or multiple occasions. Chapters 5 and 6, exploratory factor analysis, have also been rearranged, so that Chapter 5 covers a few basic factor extraction and rotation methods, for the benefit of instructors who prefer a briefer brush with EFA, and Chapter 6 treats more advanced matters. Chapter 7 has become less of a grab bag of specialized topics, with some of these (e.g., models with means, nonlinear models, and higher-order factors) being promoted to appropriate earlier chapters, and others (e.g., phantom variables) moving to an appendix. The detailed description of most goodness-of-fit indices is now in an appendix for reference rather than encumbering the main text. A few items, such as the centroid method and multivariate path models, have disappeared from the book altogether, and a few items have been added, such as sections on missing data, nonnormality, mediation, factorial invariance, and automating the construction of path diagrams. To save students labor in typing, a CD is supplied containing the various correlation and covariance matrices used in the exercises (details are given at the end of Chapter 2). A few new easy exercises have been added in the early chapters, and a number of the existing exercises have moved or changed in conformity with the text shifts. Overall, there has been a substantial expansion and updating of the reference list and the end-of-chapter notes.

I continue to be grateful to the people mentioned previously, as well as to several additional anonymous referees, and to the folks at Erlbaum: Debra Riegert has been very helpful as editor, Art Lizza continues as an invaluable resource on the production side, and of course Larry Erlbaum beams benevolently upon us all.

If you happen to notice any errors that have slipped by, I would be grateful if you would call them to my attention: loehlin@psy.utexas.edu. Enjoy the book.

**Note to the fifth edition:** Although the fifth edition adds a co-author, the book's approach remains the same as that of previous editions, as does much of the content, although we have reorganized some of it and added some as well. In addition to expanding the reference list, internet sources, and the end-of-chapter notes, there are a few other changes that apply throughout the book. First, we added discussion of packages from the R statistical program at several points in the text. While we realize not everyone uses R, we chose to include (and in some places emphasize) the program because it is free, open source, and available for almost all computer operating systems. Second, we added an Extended Example at the end of Chapters 1–6, usually using R, where we work through one of the chapter's examples in more detail to aid readers in conducting similar analyses with their own data. Third, we have removed the data CD and, instead, created an accompanying website

([routledge.com/cw/loehlin](http://routledge.com/cw/loehlin)) that contains both data and syntax for replicating the book's examples. Fourth, we re-typeset the entire book and re-created all the figures using  $\LaTeX$ . The figures are also available on the book's website.

Changes to the individual chapters are as follows. In Chapter 3, we added a section on composite reliability, and in Chapter 4 we added a section about latent variable models with categorical data. In Chapter 5 we added a section on the minimum average partial procedure. In Chapter 6 we expanded the section on hierarchical factor models to discuss bi-factor models in more depth, including a bi-factor analytical rotation. In Chapter 7, we added a new section on advanced topics. We added one appendix (H), where we provide some suggestions for communicating about latent variable models. We removed the appendix with Thurstone's Box data (now on the accompanying website), and the appendix on phantom variables has been replaced by a Chapter 3 note.

We continue to be grateful to the people mentioned previously, as well as to several anonymous referees, and to all the individuals at Routledge who have aided in this book's publication. In addition, Alex would like to thank William, Lela, Byron, Susana, Aleisa, Susanna, and especially Christine for their perpetual help and support.

If you notice any errors that have slipped by us, we would be grateful if you would bring them to our attention: [loehlin@utexas.edu](mailto:loehlin@utexas.edu) or [Alex\\_Beaujean@baylor.edu](mailto:Alex_Beaujean@baylor.edu). We hope you enjoy the book.

This page intentionally left blank

This page intentionally left blank

# Chapter 1:

## Path Models in Factor, Path, and Structural Equation Analysis

Scientists dealing with behavior, especially those who observe it occurring in its natural settings, rarely have the luxury of the simple bivariate experiment, in which a single independent variable is manipulated and the consequences observed for a single dependent variable. Even those scientists who think they do are often mistaken: The variables they directly manipulate and observe are typically not the ones of real theoretical interest but are merely some convenient variables acting as proxies or indexes for them. A full experimental analysis would again turn out to be multivariate, with a number of alternative experimental manipulations on the one side, and a number of alternative response measures on the other.

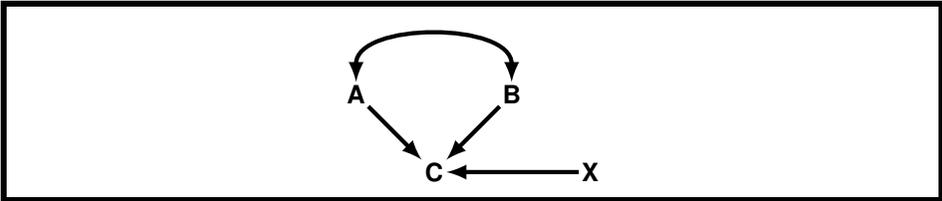
Over many years, numerous statistical techniques have been developed for dealing with situations in which multiple variables, some unobserved, are involved. Such techniques often involve large amounts of computation. Until the advent of powerful digital computers and associated software, the use of these methods tended to be restricted to the dedicated few. But in the last few decades it has been feasible for any interested behavioral scientists to take a multivariate approach to their data. Many have done so. The explosive growth in the use of computer software packages such as SPSS, SAS, and R is one evidence of this.

The common features of the methods discussed in this book are that (a) multiple variables—three or more—are involved, and that (b) one or more of these variables is unobserved, or latent. Neither of these criteria provides a decisive boundary. Bivariate methods may often be regarded as special cases of multivariate methods. Some of the methods we discuss can be—and often are—applied in situations where all the variables are, in fact, observed. Nevertheless, the main focus of our interest is on what we call, following Bentler (1980), *latent variable analysis*, a term encompassing such specific methods as factor analysis, path analysis, and structural equation modeling (SEM), all of which share these defining features.

## Path Diagrams

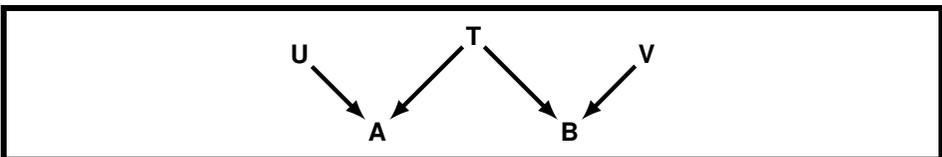
An easy and convenient representation of the relationships among a number of variables is the *path diagram*. In such a diagram we use capital letters, A, B, X, Y, and so on, to represent variables. The connections among variables are represented in path diagrams by two kinds of arrows: a straight, one-headed arrow represents a causal relationship between two variables, and a curved two-headed arrow represents a simple correlation between them.

Fig. 1.1 shows an example of a path diagram. Variables A, B, and X all are assumed to have causal effects on variable C. Variables A and B are assumed to be correlated with each other. Variable X is assumed to affect C but to be uncorrelated with either A or B. Variable C might (for example) represent young children's intelligence. Variables A and B could represent father's and mother's intelligence, assumed to have a causal influence on their child's intelligence. (The diagram is silent as to whether this influence is environmental, genetic, or both.) The curved arrow between A and B allows for the likely possibility that father's and mother's intelligence will be correlated. Arrow X represents the fact that there are other variables, independent of mother's and father's intelligence, that can affect a child's intelligence.

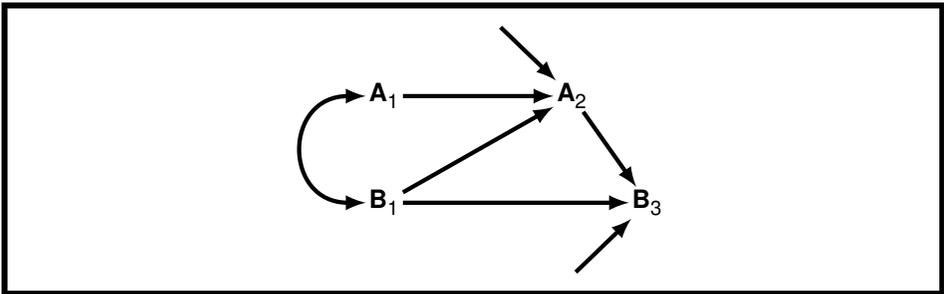


**Figure 1.1** Example of a simple path diagram.

Fig. 1.2 shows another example of a path diagram. T is assumed to affect both A and B, and each of the latter variables is also affected by an additional variable; these are labeled U and V, respectively. This path diagram could represent the reliability of a test, as described in classical psychometric test theory. A and B would stand (say) for scores on two alternate forms of a test. T would represent the unobserved true score on the trait being measured, which is assumed to affect the observed scores on both forms of the test. U and V would represent factors specific to each form of the test or to the occasions on which it was administered, which would affect any given



**Figure 1.2** Another path diagram: test reliability.



**Figure 1.3** A path diagram involving events over time.

performance but be unrelated to the true trait. (In classical psychometric test theory, the variance in A and B resulting from the influence of T is called *true score variance*, and that caused by U or V is called *error variance*. The proportion of the variance of A or B due to T is called the *reliability* of the test.)

Fig. 1.3 shows a path representation of events over time. In this case, the capital letters A and B are used to designate two variables, with subscripts to identify the occasions on which they are measured: Both A and B are measured at time 1, A is measured again at time 2, and B at time 3. In this case, the diagram indicates that both A<sub>1</sub> and B<sub>1</sub> are assumed to affect A<sub>2</sub>, but that the effect of A<sub>1</sub> on B at time 3 is wholly via A<sub>2</sub>—there is no direct arrow drawn leading from A<sub>1</sub> to B<sub>3</sub>. It is assumed that A<sub>1</sub> and B<sub>1</sub> are correlated, and that A<sub>2</sub> and B<sub>3</sub> are subject to additional influences independent of A and B, here represented by short, unlabeled arrows. These additional influences could have been labeled, say, X and Y, but are often left unlabeled in path diagrams, as here, to indicate that they refer to other, unspecified influences on the variable to which they point. Such arrows are called *residual arrows* to indicate that they represent causes residual to those explicitly identified in the diagram.

### The meaning of “cause” in a path diagram

Straight arrows in path diagrams are said to represent causal relationships—but in what sense of the sometimes slippery word “cause”? In fact, we do not need to adopt any strict or narrow definition of cause in this book, because path diagrams can be—and are—used to represent causes of various kinds, as the examples we have considered suggest. The essential feature for the use of a causal arrow in a path diagram is the assumption that a change in the variable at the tail of the arrow will result in a change in the variable at the head of the arrow, all else being equal (i.e., with all other variables in the diagram held constant). Note the one-way nature of this process—imposing a change on the variable at the head of the arrow does *not* bring about a change in the tail variable. A variety of common uses of the word “cause” can

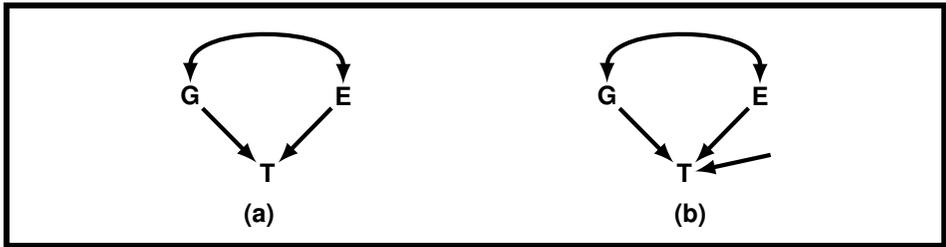
be expressed in these terms, and hence can legitimately be represented by a causal arrow in a path diagram.

### Completeness of a path diagram

Variables in a path diagram may be grouped in two classes: those that do not receive causal inputs from any other variable in the path diagram, and those that receive one or more such causal inputs. Variables in the first of these two classes are referred to as *exogenous, independent, or source* variables. Variables in the second class are called *endogenous, dependent, or downstream* variables. *Exogenous variables* (Greek: “of external origin”) are so called because their causal sources lie external to the path diagram; they are causally independent with respect to other variables in the diagram—straight arrows may lead away from them but never toward them. These variables represent causal sources in the diagram. Examples of such source variables in Fig. 1.3 are  $A_1$ ,  $B_1$ , and the two unlabeled residual variables. *Endogenous variables* (“of internal origin”) have at least some causal sources that lie within the path diagram; these variables are causally dependent on other variables—one or more straight arrows lead into them. Such variables lie causally *downstream* from source variables. Examples of downstream variables in Fig. 1.3 are  $A_2$  and  $B_3$ . In Fig. 1.2, U, T, and V are source variables, and A and B are downstream variables. Look back at Fig. 1.1. Which are the source and downstream variables in this path diagram? (We hope you identified A, B, and X as source variables, and C as downstream.)

In a proper and complete path diagram, all the source variables are interconnected by curved arrows, to indicate that they may be correlated—unless it is explicitly assumed that their correlation is zero, in which case the curved arrow is omitted. Thus the absence of a curved arrow between two source variables in a path diagram, as between X and A in Fig. 1.1, or T and U in Fig. 1.2, is not an expression of ignorance but an explicit statement about assumptions underlying the diagram.

Downstream variables, on the other hand, are never connected by curved arrows in path diagrams. (Actually, some authors use downstream curved arrows as a shorthand to indicate correlations among downstream variables caused by other variables than those included in the diagram: We use correlations between residual arrows for this purpose, which is consistent with our convention because the latter are source variables.) Residual arrows point at downstream variables, never at source variables. Completeness of a path diagram requires that a residual arrow be attached to every downstream variable unless it is explicitly assumed that all the causes of variation of that variable are included among the variables upstream from it in the diagram. (This convention is also not universally adhered to: Occasionally, path diagrams are published with the notation “residual arrows omitted.” This is an unfortunate practice



**Figure 1.4** Path diagrams illustrating the implication of an omitted residual arrow.

because it leads to ambiguity in interpreting the diagram: Does the author intend that all the variation in a downstream variable is accounted for within the diagram, or not?)

Fig. 1.4 shows an example in which the presence or absence of a residual arrow makes a difference. The source variables G and E refer to the genetic and environmental influences on a trait T. The downstream variable T in Fig. 1.4(a) has no residual arrow. That represents the assumption that the variation of T is completely explained by the genetic and environmental influences upon it. This is a theoretical assumption that one might sometimes wish to make. Fig. 1.4(b), however, represents the assumption that genetic and environmental influences are not sufficient to explain the variation of T—some additional factor or factors, perhaps measurement error or gene-environment interaction—may need to be taken into account in explaining T. Obviously, the assumptions in Figs. 1.4(a) and 1.4(b) are quite different, and one would not want it assumed that (a) was the case when in fact (b) was intended.

Finally, all significant direct causal connections between source and downstream variables, or between one downstream variable and another, should be included as straight arrows in the diagram. Omission of an arrow between  $A_1$  and  $B_3$  in Fig. 1.3 is a positive statement: that  $A_1$  is assumed to affect  $B_3$  only by way of  $A_2$ .

The notion of completeness in path diagrams should not be taken to mean that the ideal path diagram is one containing as many variables as possible connected by as many arrows as possible. Exactly the opposite is true. The smallest number of variables connected by the smallest number of arrows that can do the job is the path diagram to be sought for, because it represents the most parsimonious explanation of the phenomenon under consideration. Big, messy path diagrams are likely to give trouble in many ways. Nevertheless, often the simplest explanation of an interesting behavioral or biological phenomenon does involve causal relationships among a number of variables, not all observable. A path diagram provides a way of representing in a clear and straightforward fashion what is assumed to be going on in such a case.

Notice that most path diagrams could in principle be extended indefinitely back past their source variables: These could be taken as downstream variables in an extended path diagram, and the correlations among them explained by the linkages among their own causes. Thus, the parents in Fig. 1.1 could be taken as children in