Eur J Epidemiol
DOI 10.1007/s10654-016-0189-8

BOOK REVIEW

Tyler J VanderWeele. Explanation in causal inference: methods for mediation and interaction

Oxford University Press, 2015

Ashley I. Naimi[1]

Received: 27 July 2016 / Accepted: 30 July 2016

The field of epidemiology is primarily concerned with evaluating cause-effect relationships. Research questions are often first articulated informally using (e.g.) English language sentences. For example, "what is the effect of smoking on cardiovascular disease risk, irrespective of smoking's effect on body weight?" To answer this question, data are collected and entered into a computer, which performs a number of mathematical operations with the data to yield a number, usually interpreted as the answer to the research question of interest.

Robins long ago noted a problem with the relation between the (e.g.) English language research questions asked by epidemiologists, and the computer calculations that yield results interpreted as cause-effect relations [1]. On the one hand, the computer calculations are based on rigorously defined mathematical objects. On the other hand, the English language sentences conveying cause-effect relations are often ambiguous. Causal inference was developed to provide a formal link between the English language questions researchers ask, and the mathematical operations employed to answer them. Much of the work in causal inference is contained in technical journals, out of reach to many applied epidemiologists. *Explanation in Causal Inference: Methods for Mediation and Interaction,* by Tyler J. VanderWeele does much to resolve this obstacle [2].

Totaling 16 chapters, the book is divided into three parts: mediation, interaction, and a synthesis of mediation and interaction. The book is framed largely as a review of a vast and technical literature, with relatively little new material. The book's intended purpose is to make this literature more accessible to a wide array of applied scientists. It does so by touching upon causal inference from three different vantage points: practical, technical, and philosophical.

The practical dimension of this book is by far its major strength. The main text avoids heavy technical notation (the appendices consist almost entirely of rigorous mathematical proofs that underlie the methods outlined in the book). It is replete with detailed applied examples of mediation and interaction analyses that can be used as a template for researchers interested in asking similar questions in other substantive areas. Peppered throughout the book are useful tips on practical questions, such as when to include an exposure-mediator interaction (p45), how to present interaction analysis results (p270), and welcome guidance on conducting sensitivity/bias analyses in studies using mediation and interaction analysis. It also has a heavy software component that illustrates how SAS, Stata, R, and SPSS can be used to conduct mediation and interaction analysis, with code embedded directly in the text.

This book's technical component is contained within the roughly 200 Appendix pages, providing a formal justification for nearly the entirety of the book. The appendix starts with a brief introduction to potential outcomes, and then quickly proceeds to articulate the formal conditions for identifying direct and indirect effects. After these initial segments, the appendix then systematically follows each section of the main text, providing proofs and technical details of nearly each of the procedures outlined in the book. This correspondence between the main text and technical appendix is a resource for those interested in such details, as it is not difficult to find the set of equations that

✉ Ashley I. Naimi
ashley.naimi@pitt.edu

[1] Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, 130 DeSoto Street, Parran 503, Pittsburgh, PA 15261, USA

Published online: 12 August 2016

formalize a given procedure in just about any part of the book. Additionally, in a few instances, the formal proofs are accompanied by a short explanation of each step in the proof. Thus, in some instances, mathematically inclined epidemiologists may be able to follow the specific line of reasoning used to justify some of the methods for mediation or interaction, and thus develop a deeper understanding of the logic behind certain epidemiologic methods.

Philosophically, this book covers several subjects related to causal inference, mostly from the perspective of traditional Western philosophy. In particular, the book's concluding chapter (chapter 16) touches upon issues related to natural laws, human agency, and arguments for a "first cause" (i.e., the cosmological argument). In it, VanderWeele outlines the sweeping intellectual vistas that have long been devoted to understanding more profound existential questions about where it all came from, and why. It is a welcome initiation to topics that one might surmise are unfamiliar to most epidemiologists.

Practically, this book will be an invaluable asset for students and researchers interested in developing a more complete and nuanced understanding of the analytic tools routinely used in epidemiology, as well as novel approaches that are only beginning to appear, particularly for questions related to mediation and interaction. Given the rapid progress that has been made in these areas, this book is a much-needed introduction, and fills an important gap between more classical and contemporary epidemiologic methods. However, as an introduction, it does not cover more advanced (and arguably more robust) analytic techniques that are at the vanguard of methodological development in causal inference. Among these include targeted maximum likelihood methods, the role of data mining and machine learning algorithms in evaluating causation, and other advances semiparametric theory. Additionally, it does not provide an in depth and intuitive treatment of the fundamentals of causal inference—a task to be covered by future texts. Nevertheless, this book can serve as an important stepping-stone to both.

Overall, VanderWeele has masterfully penned a widely accessible treatise on how to think rigorously about data analysis in epidemiology. Of this book, one can safely state that there are few others like it in the field.

References

1. Robins JM. Addendum to "a new approach to causal inference in mortality studies with a sustained exposure period-application to control of the healthy worker survivor effect". Math Model. 1987;14(9–12):923–45.
2. VanderWeele TJ. Explanation in causal inference: methods for mediation and interaction. Oxford: Oxford University Press; 2015.