

NGHIÊN CỨU KHOA HỌC XÃ HỘI: NGUYÊN TẮC, PHƯƠNG PHÁP VÀ THỰC HÀNH



ANOL BHATTACHERJEE

PHAN VIẾT PHONG

CAO NGỌC ANH

**SOCIAL SCIENCE RESEARCH:
PRINCIPLES, METHODS, AND PRACTICES**

Anol Bhattacharjee, Ph.D.

University of South Florida

Tampa, Florida, USA

abhatt@usf.edu

Second Edition

Copyright © 2012 by Anol Bhattacharjee

Published under the

Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License



Dịch từ nguyên bản tiếng Anh:

Social Science Research: Principles, Methods, and Practices, 2nd edition

By Anol Bhattacharjee

First published 2012

ISBN-13: 978-1475146127

ISBN-10: 1475146124

Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License:

Users are free to use, copy, share, distribute, display, and reference this book under the following conditions:

- **ATTRIBUTION:** Whole or partial use of this book should be attributed (referenced or cited) according to standard academic practices.
- **NON-COMMERCIAL USE:** This book may not be used for commercial purposes.
- **SHARE ALIKE:** Users may, transform, or build upon this book, but must distribute the resulting work under the same or similar license as this one.

For any reuse or distribution, the license terms of this work must be clearly specified. Your fair use and other rights are in no way affected by the above.

Copyright © 2012 by Anol Bhattacharjee.

Copyright Vietnamese version 2015 by Phan Viet Phong and Cao Ngoc Anh.

MỤC LỤC

Lời nói đầu	3
Giới thiệu về nghiên cứu	
Chương 1 Khoa học và nghiên cứu khoa học	7
Chương 2 Tư duy của nhà nghiên cứu	20
Chương 3 Tiến trình nghiên cứu	31
Chương 4 Lý thuyết trong nghiên cứu khoa học	43
Cơ sở của nghiên cứu thực nghiệm	
Chương 5 Thiết kế nghiên cứu	58
Chương 6 Phương thức đánh giá các phạm trù	69
Chương 7 Giá trị và độ tin cậy	86
Chương 8 Chọn mẫu	100
Thu thập dữ liệu	
Chương 9 Nghiên cứu khảo sát	112
Chương 10 Nghiên cứu thực nghiệm	127
Chương 11 Nghiên cứu trường hợp	141
Chương 12 Nghiên cứu điển giải	156
Phân tích dữ liệu	
Chương 13 Phân tích định tính	169
Chương 14 Phân tích định lượng: Thống kê mô tả	176
Chương 15 Phân tích định lượng: Thống kê suy diễn	188
Kết luận	
Chương 16 Đạo đức nghiên cứu	199
Phụ lục	209

LỜI NÓI ĐẦU

Cuốn sách này nhằm mục đích giới thiệu cho học viên cao học và nghiên cứu sinh về cách tiến hành nghiên cứu khoa học trong các ngành khoa học xã hội, kinh doanh, giáo dục, y tế cộng đồng và các chuyên ngành khác có liên quan. Sách được biên soạn dựa trên các bài giảng của cá nhân trong hơn một thập kỷ giảng dạy phương pháp nghiên cứu khoa học cho các khoá nghiên cứu sinh tại trường Đại học South Florida. Mục đích của cuốn sách nhằm hướng tới độc giả là các nghiên cứu sinh tiến sĩ, học viên cao học, các nhà nghiên cứu trẻ và các giáo sư giảng dạy bộ môn phương pháp nghiên cứu khoa học. Đương nhiên, những nhà nghiên cứu khác cũng có thể sử dụng tham khảo như một tài liệu nhỏ gọn và tiện dụng.

Câu hỏi đầu tiên và quan trọng nhất mà độc giả thường quan tâm là liệu có sự khác biệt nào giữa cuốn sách này với các cuốn sách khác viết về phương pháp nghiên cứu khoa học đang bán trên thị trường? Xin trả lời rằng có bốn sự khác biệt cơ bản. *Thứ nhất*, không giống như các sách khác, cuốn sách này không chỉ viết về “phương pháp nghiên cứu” (thu thập và phân tích dữ liệu thực nghiệm), mà còn đề cập toàn bộ “quá trình nghiên cứu” từ lúc bắt đầu đến khi kết thúc. Phương pháp nghiên cứu chỉ là một giai đoạn trong quá trình nghiên cứu và có lẽ là một trong những công việc đơn giản và rõ ràng nhất. Hầu hết các sách hiện có đi sâu giới thiệu, phân tích về các phương pháp nghiên cứu, nhưng lại bỏ sót các vấn đề phức tạp, quan trọng hơn và khó hình dung về cấu trúc, như lý thuyết hóa và tư duy khoa học của nhà nghiên cứu. Những vấn đề này lại thường là điều kiện tiên quyết cho thành công của các nghiên cứu thực nghiệm. Theo kinh nghiệm cá nhân, hầu hết nghiên cứu sinh tiến sĩ sử dụng khá thành thạo các phương pháp nghiên cứu, nhưng nhiều học viên còn gặp lúng túng trong việc phát hiện câu hỏi nghiên cứu thú vị, hữu ích hoặc xây dựng các lý thuyết khoa học. Để khắc phục những khó khăn này, tôi có dành nhiều chương cho chủ đề đó như: “Tư duy của nhà nghiên cứu” và “Lý thuyết trong nghiên cứu khoa học”. Đây là những kỹ năng cơ bản cho những người mới bắt đầu bước vào nghiên cứu khoa học.

Thứ hai, cuốn sách được thiết kế một cách ngắn gọn và xúc tích. Khi biên soạn, tôi quyết định chỉ đi vào phân tích những khái niệm cơ bản nhất để cho độc giả tập trung vào những nội dung thiết yếu mà không đề cập đến những nội dung bên lề, ít liên quan. Hầu

hết chuyên đề nghiên cứu tiến sĩ có kế thừa khá nhiều tài liệu tham khảo liên quan và cuốn sách này sẽ đề cập tất cả tài liệu đó bằng cách tóm tắt các khái niệm quan trọng một cách cô đọng. Nội dung cuốn sách không đưa vào khối lượng lớn các tài liệu có liên quan nhằm tránh gây áp lực cho sinh viên.

Thứ ba, độc giả có thể tải miễn phí cuốn sách này trên mạng, không chỉ phiên bản hiện tại mà các lần chỉnh lý tiếp theo. Bạn có thể đọc sách này trên các thiết bị điện tử như Kindle e-Book, Apple iBook hiện có giá rất phải chăng. Nhiều người hỏi tôi tại sao lại cung cấp tài liệu miễn phí trong khi tôi hoàn toàn có thể bán chúng để lấy tiền. Câu trả lời là tôi muốn giúp sinh viên bớt áp lực về giá cả đắt đỏ của nhiều giáo trình và sách tham khảo, mà còn vì tôi cho rằng, tri thức khoa học không nên bị hạn chế bởi những rào cản tiếp cận kiểu như giá cả và việc mua bán. Tiến bộ khoa học chỉ có thể đạt được nếu như sinh viên và các học giả trên toàn thế giới có khả năng tiếp cận thuận lợi nhất tới khoa học và cuốn sách này là đóng góp nhỏ bé của tôi cho sự nghiệp chung của khoa học. Tuy nhiên, miễn phí không có nghĩa là “chất lượng kém hơn”. Một vài thứ quan trọng nhất cho cuộc sống như không khí, nước và ánh sáng mặt trời là miễn phí. Rất nhiều nguồn thông tin trên Google cũng miễn phí - điều ai cũng có thể hình dung rõ ràng là chúng ta đang sống trong thời đại Internet không thể thiếu Google. Một vài chương trình phần mềm phức tạp nhất hiện nay như Linux và Apache cũng miễn phí và cuốn sách này không phải là ngoại lệ.

Thứ tư, tôi muốn các phiên bản dịch thuật sang các ngôn ngữ khác của cuốn sách cũng hoàn toàn miễn phí cho các khoá học hiện thời. Bởi vậy, tôi hứa dịch cuốn sách sang tiếng Trung, Pháp, Indonesia, Hàn Quốc, Bồ Đào Nha, Tây Ban Nha (hy vọng có thể đưa vào sử dụng năm 2012). Tôi cũng thật sự mong đợi các nhà nghiên cứu, các giáo sư khác có khả năng dịch cuốn sách này sang tiếng A-rập, Đức và các ngôn ngữ khác khi có nhu cầu. Nếu bạn là dịch giả của cuốn sách này trong tương lai, xin hãy lưu ý rằng sẽ không có lợi nhuận tài chính hay bất kỳ tiền nhuận bút nào cho công việc dịch thuật của bạn, bởi cuốn sách này phải được miễn phí. Tuy vậy, tôi sẽ rất hạnh phúc xem bạn như là đồng tác giả trong bản dịch ngôn ngữ bản địa.

Cuốn sách được cấu trúc thành 16 chương cho một học kỳ 16 tuần. Các giáo sư hoặc giảng viên có thể bổ sung, rút ngắn, kéo dài hoặc điều chỉnh sao cho phù hợp với nhu cầu cụ thể trong chương trình giảng dạy. Chẳng hạn, tôi không giảng Chương 14 và 15 trong các lớp học của tôi, bởi chúng tôi đã nghiên cứu về khoa học thống kê và đề cập nội dung này kỹ hơn rồi. Thay vào đó, tôi sử dụng 2 tuần này vào nghiên cứu các lý thuyết (Chương 3), trong đó một tuần cho thảo luận và nghiên cứu các bài báo khoa học (không có trong

sách) và một tuần cho thi cuối kỳ. Tuy nhiên, việc nghiên cứu Chương 14 và 15 là cần thiết cho các chương trình học không có phần về phân tích thống kê trong nghiên cứu. Một đề cương khoá học tôi sử dụng những bài giảng về chương trình đào tạo tiến sĩ kinh doanh được cung cấp ở phần phụ lục.

Cuối cùng, tôi có kế hoạch liên tục cập nhật cuốn sách này dựa trên các xu hướng đang xuất hiện trong nghiên cứu khoa học. Nếu như có bất kỳ nội dung mới và cần thiết nào mà bạn muốn đọc trong những lần xuất bản tiếp theo, xin hãy thông tin cho tôi, tôi sẽ cố gắng hết sức để cập nhật chúng. Tôi thật sự cảm kích và đánh giá cao các nhận xét, phê bình hay sửa chữa cho bất kỳ nội dung nào của cuốn sách.

Anol Bhattacharjee

Email: abhattach@usf.edu

Chương 1

KHOA HỌC VÀ NGHIÊN CỨU KHOA HỌC

Nghiên cứu là gì? Tùy thuộc người được hỏi, bạn sẽ nhận được câu trả lời hết sức khác nhau cho câu hỏi tương chừng rất đơn giản này. Một số người nói rằng họ thường xuyên nghiên cứu trên các websites trực tuyến để tìm kiếm nơi bán hàng hoá hoặc cung cấp dịch vụ tốt nhất mà họ mong muốn. Các kênh tin tức trên ti vi cũng tiến hành nhiều hình thức nghiên cứu ý kiến của khán giả về các chủ đề mà công chúng quan tâm như cuộc bầu cử sắp đến hoặc các dự án của chính phủ. Sinh viên đại học nghiên cứu trên Internet các thông tin cần thiết để hoàn thành luận văn, đồ án hoặc các bài báo khoa học đến hạn hoàn thành. Sinh viên ra trường đang làm việc trong các dự án của giáo sư có thể nghiên cứu thu thập hoặc phân tích dữ liệu liên quan đến dự án. Các doanh nghiệp và công ty tư vấn nghiên cứu tìm kiếm các giải pháp để khắc phục những vấn đề về tổ chức như rào cản chuỗi cung ứng hoặc nhận diện mô hình mua sắm của khách hàng. Tuy nhiên, các hình thức nghiên cứu trên không thể được coi là “nghiên cứu khoa học” trừ khi: (1) nghiên cứu đó có đóng góp mới cho khoa học, và (2) được tiến hành bằng phương pháp khoa học. Chương này sẽ khảo cứu các thuật ngữ trên.

Khoa học

Khoa học là gì? Đối với một số người, khoa học liên quan đến các khóa học bậc cao đẳng hoặc đại học, chẳng hạn như vật lý, hóa học và sinh học - có nghĩa là chỉ dành cho các học sinh xuất sắc. Đối với những người khác, khoa học là một nghề do các nhà khoa học khoác áo trắng thực hiện bằng cách sử dụng thiết bị chuyên ngành trong các phòng thí nghiệm. Theo từ nguyên, "*khoa học*" có nguồn gốc từ chữ Latin *scientia* có nghĩa là tri thức. Khoa học là hệ thống tri thức được tổ chức theo các lĩnh vực và đòi hỏi sử dụng "phương pháp khoa học" (phương pháp khoa học được trình bày dưới đây). Khoa học có thể được phân thành hai nhóm lớn: khoa học tự nhiên và khoa học xã hội.

Khoa học tự nhiên là khoa học nghiên cứu các đối tượng hoặc hiện tượng tự nhiên, chẳng hạn như ánh sáng, vật chất, trái đất, các thiên thể hoặc cơ thể con người. Khoa học tự nhiên lại có thể được phân loại tiếp thành các khoa học vật chất, khoa học trái đất, khoa học sự sống và các khoa học khác. Khoa học vật chất bao gồm các bộ môn khoa học như

vật lý (khoa học của các đối tượng vật lý), hóa học (khoa học vật chất) và thiên văn học (khoa học của các đối tượng thiên thể). Khoa học trái đất bao gồm các bộ môn khoa học như địa chất học (khoa học của trái đất). Khoa học sự sống bao gồm các bộ môn như sinh học (khoa học của cơ thể con người) và thực vật học (khoa học của thực vật).

Khoa học xã hội là khoa học nghiên cứu con người hoặc cộng đồng người chẳng hạn như các nhóm xã hội, doanh nghiệp, hiệp hội hoặc kinh tế và hành vi cá nhân, tập thể. Khoa học xã hội có thể được phân loại thành các bộ môn khoa học như tâm lý học (khoa học về hành vi con người), xã hội học (khoa học về các nhóm xã hội) và kinh tế học (khoa học của các doanh nghiệp, thị trường và kinh tế).

Khoa học tự nhiên và khoa học xã hội có sự khác biệt trên một vài phương diện. Khoa học tự nhiên đòi hỏi sự chính xác nghiêm ngặt, rõ ràng và không phụ thuộc vào người tiến hành các nghiên cứu khoa học. Ví dụ như trong vật lý học, thí nghiệm đo tốc độ lan truyền âm thanh qua một môi trường truyền dẫn hoặc đo chỉ số khúc xạ của nước, thì kết quả thí nghiệm thu được luôn giống nhau, không phân biệt thời gian hoặc địa điểm thí nghiệm hoặc người tiến hành thí nghiệm. Nếu hai sinh viên cùng làm một thí nghiệm vật lý mà nhận được hai giá trị có đặc tính vật lý khác nhau, thì có nghĩa là một hoặc cả hai sinh viên đó mắc lỗi. Tuy nhiên, đối với khoa học xã hội thì không thể kết luận như vậy, bởi lẽ khoa học xã hội yêu cầu ít hơn về sự chính xác, cụ thể và rõ ràng. Ví dụ, nếu bạn đo lường mức độ hạnh phúc của một người bằng cách sử dụng công cụ giả thuyết, bạn có thể nhận ra rằng người đó cảm thấy hạnh phúc hay bất hạnh, buồn chán ở những ngày khác nhau và đôi khi ở những thời điểm khác nhau trong cùng ngày. Hạnh phúc của một người tùy thuộc vào thông tin người đó nhận được trong ngày hoặc các sự kiện diễn ra hôm trước. Hơn nữa, không có một công cụ hoặc chỉ số nào có thể đo lường chính xác hạnh phúc của một người. Vì thế trong cùng thời điểm, một công cụ này có thể xác định người này là “hạnh phúc hơn” thì một công cụ đo thứ hai có thể cho ra kết quả ngược lại rằng người đó “kém hạnh phúc”.

Nói cách khác, tồn tại mức độ khác biệt lớn về đo lường trong khoa học xã hội cũng như sự thiếu tin cậy và ít sự đồng thuận về các kết luận trong khoa học xã hội. Trong khi bạn sẽ không tìm thấy sự bất đồng giữa các nhà khoa học tự nhiên về tốc độ của ánh sáng hay tốc độ của trái đất quay xung quanh mặt trời, nhưng bạn sẽ nhận thấy sự khác biệt rất lớn giữa các nhà khoa học xã hội về cách giải quyết một vấn đề xã hội, như ngăn ngừa khủng bố quốc tế, vực dậy nền kinh tế khỏi sự suy thoái. Bất kỳ sinh viên nghiên cứu khoa

học xã hội cũng phải nhận thức đầy đủ về sự lý giải còn mơ hồ, thiếu chắc chắn lẫn sai sót trong khoa học, phản ánh sự biến thiên cao của các khách thể nghiên cứu xã hội.

Khoa học cũng có thể được phân loại dựa trên mục đích nghiên cứu. **Khoa học cơ bản** (basic sciences) còn gọi là khoa học thuần túy, là khoa học giải thích bản chất sự vật, các mối quan hệ tương tác và quy luật phổ biến của sự vật. Ví dụ như vật lý học, toán học và sinh học. **Khoa học ứng dụng** (applied sciences) còn được gọi là khoa học thực hành, là khoa học áp dụng tri thức từ khoa học cơ bản vào thực tế. Ví dụ, kỹ thuật là một khoa học ứng dụng các định luật vật lý và hóa học vào thực tiễn như xây dựng các cây cầu chịu được tải trọng lớn hơn hoặc chế tạo động cơ sử dụng nhiên liệu hiệu quả hơn; hoặc như y học là một khoa học ứng dụng các quy luật sinh học để chữa bệnh. Cả khoa học cơ bản và khoa học ứng dụng đều là cần thiết cho sự phát triển của nhân loại. Tuy nhiên, khoa học ứng dụng không thể đứng riêng rẽ mà phải dựa vào khoa học cơ bản trong mỗi bước phát triển. Tất nhiên, các tập đoàn công nghiệp và doanh nghiệp tư nhân thường thiên về khoa học ứng dụng nhằm mang lại giá trị thực tế cho họ, trong khi đó, các trường đại học coi trọng cả khoa học cơ bản và khoa học ứng dụng.

Tri thức khoa học

Mục đích của khoa học là tạo ra tri thức khoa học. **Tri thức khoa học** (scientific knowledge) là tổng hợp các tri thức về các quy luật, lý thuyết giải thích hiện tượng hoặc trạng thái sự vật bằng cách sử dụng phương pháp khoa học. Quy luật hay định luật (laws) trạng thái tồn tại của các hiện tượng hoặc hành vi; còn lý thuyết (theories) là giải thích một cách hệ thống cơ chế tồn tại của các hiện tượng hay hành vi đó. Ví dụ trong vật lý, định luật Newton về chuyển động mô tả điều xảy ra khi một vật trong trạng thái đứng yên hoặc chuyển động (Định luật thứ nhất của Newton), cần có lực tác động nào để tạo ra chuyển động cho vật đang đứng yên hoặc làm cho vật đang chuyển động sang trạng thái đứng yên (Định luật thứ hai của Newton), mô tả và giải thích điều xảy ra khi hai vật va chạm với nhau (Định luật thứ ba của Newton). Ba định luật này tạo thành nền tảng cho cơ học cổ điển - *lý thuyết về sự chuyển động của vật*. Tương tự, lý thuyết quang học giải thích tính chất của ánh sáng và sự lan truyền ánh sáng trong các môi trường khác nhau; lý thuyết điện từ cắt nghĩa tính chất của điện và nguyên lý tạo ra điện; cơ học lượng tử giải thích các thuộc tính của hạt trong nguyên tử; nhiệt động học giải thích đặc tính của năng lượng và vận hành cơ khí. Các khóa học vật lý bậc đại học thường có các chương dành riêng cho từng lý thuyết trên.

Các lý thuyết trong các khoa học xã hội cũng tương tự như vậy. Chẳng hạn, lý thuyết thiên lệch nhận thức (cognitive dissonance theory) trong tâm lý học giải thích sự phản ứng khác nhau của con người khi quan sát một sự kiện; lý thuyết răn đe tổng hợp (general deterrence theory) giải thích lý do một số người tham gia vào các hành vi lệch chuẩn hoặc phạm tội như hành vi tải nhạc bất hợp pháp hoặc vi phạm bản quyền phần mềm; lý thuyết hành vi hoạch định (theory of planned behavior) giải thích cách thức con người lập luận để đưa ra các lựa chọn có ý thức trong cuộc sống hàng ngày của họ.

Mục đích của nghiên cứu khoa học là khám phá, phát hiện các quy luật, định luật và thiết lập các lý thuyết tiên khởi để giải thích hiện tượng tự nhiên hoặc xã hội, hay nói cách khác là tạo ra tri thức khoa học. Điều quan trọng cần hiểu rằng tri thức này có thể chưa hoàn thiện hoặc thậm chí còn khác xa sự thật. Đôi khi không có một chân lý phổ biến mà là trạng thái có “nhiều sự thật”. Cần hiểu rằng các lý thuyết dựa trên cơ sở tri thức khoa học chỉ là sự giải thích một hiện tượng cụ thể do một nhà khoa học đưa ra. Như vậy, lý thuyết giải thích đầy đủ hay sơ lược tùy thuộc vào mức độ những giải thích đó phù hợp với thực tế đến đâu và do đó lý thuyết là hoàn thiện hay chưa hoàn thiện. Sự tiến bộ của khoa học đánh dấu bằng sự tiến triển trong quá trình hoàn thiện lý thuyết thông qua những tiến bộ trong quan sát nhờ sử dụng phương tiện chính xác hơn và suy luận logic hợp lý hơn.

Việc tìm ra quy luật hoặc lý thuyết khoa học phải thông qua quá trình suy luận logic và luận cứ. Suy luận logic (lý thuyết) và luận cứ (quan sát) là hai trụ cột tạo ra tri thức khoa học. Trong khoa học, lý thuyết và quan sát có quan hệ tương hỗ và phụ thuộc lẫn nhau. Lý thuyết cung cấp ý nghĩa và tầm quan trọng của những cái quan sát; còn quan sát giúp kiểm chứng hoặc chỉnh lý các lý thuyết đang tồn tại hoặc xây dựng lý thuyết mới. Bất kỳ cách thức khác thu thập tri thức, chẳng hạn như bằng đức tin hoặc quyền uy đều không thể được coi là khoa học.

Nghiên cứu khoa học

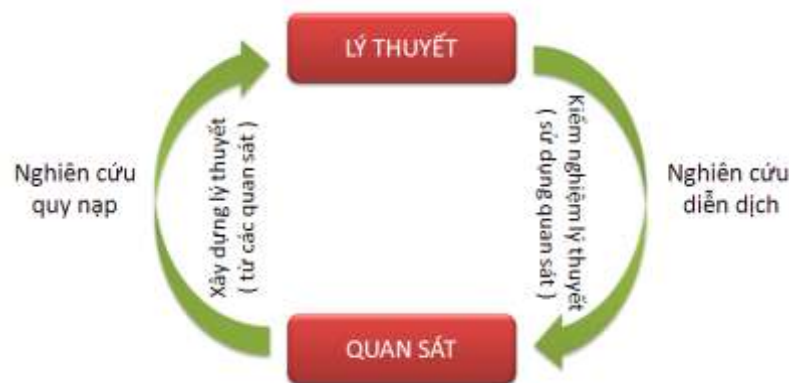
Khẳng định lý thuyết và quan sát là hai trụ cột của khoa học thì nghiên cứu khoa học được thực hiện ở hai cấp độ: *lý thuyết* và *thực nghiệm*. Cấp độ lý thuyết liên quan đến phát triển các khái niệm trừu tượng về hiện tượng tự nhiên hoặc xã hội và mối liên hệ giữa các khái niệm này (tức là xây dựng “lý thuyết”). Trong khi cấp độ thực nghiệm lại liên quan đến việc kiểm nghiệm mức độ phù hợp với thực tiễn của các khái niệm lý thuyết và mối liên hệ giữa các khái niệm để xem xét chúng phản ánh mức độ quan sát thực tế nhằm đạt mục tiêu cuối cùng là tạo ra những lý thuyết hoàn chỉnh hơn. Theo thời gian, lý thuyết

ngày càng hoàn thiện hơn (tức là phù hợp hơn với thực tế được quan sát) và khoa học tiếp tục phát triển. Nghiên cứu khoa học liên quan đến việc hoán chuyển qua lại liên tục giữa lý thuyết và quan sát thực nghiệm. Cả lý thuyết và quan sát là hai thành phần thiết yếu của nghiên cứu khoa học. Vì thế, chỉ dựa trên những quan sát để suy luận mà bỏ qua lý thuyết thì không được coi là nghiên cứu khoa học đúng nghĩa.

Tùy thuộc chuyên môn đào tạo và sở thích của người nghiên cứu, nghiên cứu khoa học có thể thực hiện bằng một trong hai hình thức: quy nạp hoặc diễn dịch. Trong nghiên cứu quy nạp (inductive research), mục tiêu của người nghiên cứu là đưa ra các khái niệm và cấu trúc lý thuyết từ các dữ liệu quan sát. Còn trong nghiên cứu diễn dịch (deductive research), mục tiêu của người nghiên cứu là để kiểm tra các khái niệm và các mô hình lý thuyết bằng cách sử dụng các dữ liệu thực nghiệm mới. Vì thế, nghiên cứu quy nạp cũng được coi là *nghiên cứu xây dựng lý thuyết (theory-building)* và nghiên cứu diễn dịch là *nghiên cứu kiểm nghiệm lý thuyết (theory-testing)*. Ở đây cần lưu ý rằng mục tiêu của kiểm nghiệm lý thuyết không phải là chỉ để kiểm nghiệm một lý thuyết, mà có thể là để chỉnh lý, nâng cao và mở rộng lý thuyết. Hình 1.1 mô tả sự bổ sung của nghiên cứu quy nạp và diễn dịch. Lưu ý rằng nghiên cứu quy nạp và diễn dịch là hai nửa của chu trình nghiên cứu liên tục lặp lại giữa lý thuyết và quan sát. Bạn không thể tiến hành nghiên cứu quy nạp hoặc diễn dịch nếu bạn không quen với cả hai thành phần lý thuyết và dữ liệu của nghiên cứu. Đương nhiên, một người nghiên cứu có kinh nghiệm có thể nghiên cứu toàn bộ theo chu trình và có thể tiến hành đồng thời cả hai nghiên cứu quy nạp và diễn dịch.

Cần chú ý rằng cả nghiên cứu xây dựng lý thuyết (nghiên cứu quy nạp) và nghiên cứu kiểm nghiệm lý thuyết (nghiên cứu diễn dịch) đều quan trọng đối với sự tiến bộ của khoa học. Lý thuyết không có giá trị nếu không phù hợp với thực tế. Tương tự như vậy, hàng núi dữ liệu cũng trở nên vô dụng cho đến khi chúng có đóng góp cho việc xây dựng một lý thuyết có ý nghĩa. Thay vì xem xét hai quá trình trong mối quan hệ vòng tròn như thể hiện trong Hình 1.1, có lẽ tốt hơn là cần xem chúng theo hình xoắn tròn ốc, theo đó mỗi chu kỳ tương tác giữa lý thuyết và dữ liệu quan sát sẽ góp phần diễn giải rõ hơn các hiện tượng và hoàn chỉnh lý thuyết. Mặc dù cả hai nghiên cứu quy nạp và diễn dịch đều rất quan trọng đối với sự tiến bộ của khoa học, nhưng dường như nghiên cứu quy nạp (xây dựng lý thuyết) có giá trị hơn trong nghiên cứu xây dựng lý thuyết tiên khởi (prior theories) hoặc nghiên cứu giải thích; trong khi nghiên cứu diễn dịch (kiểm nghiệm lý thuyết) sẽ phát huy hiệu quả hơn khi đã có nhiều lý thuyết về cùng một hiện tượng và nhà nghiên cứu muốn

tìm hiểu xem lý thuyết nào trong số này phù hợp nhất trong một môi trường, hoàn cảnh nhất định.



Hình 1.1. Chu trình nghiên cứu

Xây dựng lý thuyết và kiểm nghiệm lý thuyết là nhiệm vụ đặc biệt khó khăn trong các ngành khoa học xã hội, bởi vì tính chất trừu tượng của các khái niệm lý thuyết, sự hạn chế của các công cụ để đo lường và khả năng hiện diện của nhiều yếu tố không mong muốn ảnh hưởng đến các hiện tượng nghiên cứu. Việc bác bỏ các lý thuyết không phù hợp cũng rất khó khăn. Ví dụ, học thuyết của Karl Marx coi chủ nghĩa cộng sản là một phương tiện đảm bảo tính hiệu quả cho sự phát triển kinh tế đã đứng vững trong nhiều thập kỷ trước khi nó mất uy tín do thua kém chủ nghĩa tư bản trong việc thúc đẩy tăng trưởng kinh tế và phúc lợi xã hội. Nền kinh tế cộng sản trước đây như Liên Xô và Trung Quốc cuối cùng đã chuyển đổi sang nền kinh tế tư bản với đặc trưng là tối đa hóa lợi nhuận của doanh nghiệp tư nhân. Tuy nhiên, sự sụp đổ của việc cho vay thế chấp và lĩnh vực tài chính tại Hoa Kỳ gần đây chứng tỏ chủ nghĩa tư bản mà cũng có khuyết tật và không còn hiệu quả trong việc thúc đẩy tăng trưởng kinh tế và phúc lợi xã hội như được kỳ vọng trước đó. Không giống như trong khoa học tự nhiên, các lý thuyết khoa học xã hội rất hiếm khi hoàn hảo. Điều đó tạo cho các nhà nghiên cứu cơ hội phát triển những lý thuyết hiện có hoặc xây dựng các lý thuyết thay thế.

Thực hiện nghiên cứu khoa học đòi hỏi hai nhóm kỹ năng - lý thuyết và phương pháp luận - để tiến hành nghiên cứu ở cấp độ lý thuyết và cấp độ thực nghiệm tương ứng. Kỹ năng phương pháp luận (know-how - biết bằng cách nào) là tương đối chuẩn mực, bất biến trong tất cả các bộ môn khoa học và dễ đạt được thông qua các chương trình đào tạo tiến sĩ. Còn kỹ năng lý thuyết (know-what - biết được cái gì) thì khó khăn hơn, đòi hỏi phải có nhiều năm quan sát, suy tư và là kỹ năng ẩn mà không thể “dạy”

được mà do tích lũy qua kinh nghiệm. Tất cả các nhà khoa học vĩ đại nhất trong lịch sử nhân loại, như Galileo, Newton, Einstein, Neils Bohr, Adam Smith, Charles Darwin và Herbert Simon, là các nhà lý thuyết bậc thầy đã xây dựng các định đề dẫn đến thay đổi chiều hướng của khoa học. Kỹ năng phương pháp luận là cái cần thiết để trở thành một nhà nghiên cứu thông thường, còn kỹ năng lý thuyết là cái cần thiết để trở thành nhà nghiên cứu bậc cao.

Phương pháp khoa học

Trong phần trước, chúng ta đã coi khoa học là tri thức thu thập được thông qua phương pháp khoa học. Vậy “phương pháp khoa học” hiểu một cách chính xác là gì? **Phương pháp khoa học** (scientific method) là một tập hợp các kỹ thuật được chuẩn hóa dùng để tạo ra tri thức khoa học, chẳng hạn như cách quan sát, cách giải thích và cách khái quát kết quả đã thu nhận được. Phương pháp khoa học cho phép các nhà nghiên cứu kiểm chứng độc lập, khách quan các lý thuyết, các phát hiện trước đó và mở ra chủ đề tranh luận, sửa đổi hoặc cải tiến lý thuyết hoặc phát hiện trước. Phương pháp khoa học phải đáp ứng bốn đặc điểm sau:

- Tính lặp lại (Replicability): Những người khác sẽ có thể độc lập thực hiện lại các nghiên cứu khoa học với cùng phương pháp và đạt được kết quả giống nhau hoặc tương tự.

- Tính chính xác (Precision): Thường rất khó đo lường, song các khái niệm lý thuyết phải được định nghĩa chính xác để những người khác có thể sử dụng những định nghĩa đó để đo lường khái niệm và kiểm nghiệm lý thuyết.

- Tính phản nghiệm (Falsifiability): Một lý thuyết phải được trình bày theo cách mà nó có thể bị bác bỏ. Lý thuyết mà không thể kiểm nghiệm được hoặc không thể bị bác bỏ thì không phải là lý thuyết khoa học; bất kỳ tri thức nào tạo ra như vậy đều không phải là tri thức khoa học. Một lý thuyết xây dựng trên các thuật ngữ mơ hồ hoặc các khái niệm không chính xác, không kiểm nghiệm được thì không phải là lý thuyết khoa học. Ý tưởng của Sigmund Freud về phân tâm học thuộc loại này và do đó không được coi là một “lý thuyết”, mặc dù phân tâm học có giá trị thực tiễn trong điều trị một số loại bệnh.

- Tính tối giản (Parsimony): Khi có nhiều giải thích về một hiện tượng, các nhà khoa học chỉ luôn chấp nhận lời giải thích đơn giản nhất hoặc logic hợp lý nhất. Khái niệm này là gọi là tính tối giản hoặc “nguyên lý dao cạo Ockham” (nguyên lý Ockham’s razor). Tiêu chuẩn tối giản giúp hạn chế việc các nhà khoa học theo đuổi các lý thuyết quá phức tạp

hoặc kỳ dị với số lượng vô tận của các khái niệm và mối quan hệ chỉ để giải thích một khía cạnh nhỏ của những vấn đề không có gì đặc biệt.

Ngành học nào không chấp nhận sử dụng phương pháp khoa học để kiểm nghiệm các quy luật hoặc các lý thuyết cơ bản đều không thể được gọi là "khoa học". Chẳng hạn, thần học (nghiên cứu của tôn giáo) không phải là khoa học, bởi lẽ những ý niệm thần học (sự hiện diện của Chúa Trời) không thể kiểm chứng bằng các nhà quan sát độc lập, không thoả mãn tính lặp lại, tính chính xác, tính phản nghiệm và tối giản. Tương tự như vậy, nghệ thuật, âm nhạc, văn học, nhân văn và pháp luật cũng không được coi là khoa học, mặc dù cũng mang tính sáng tạo và cố gắng tạo ra giá trị trong từng lĩnh vực.

Khi áp dụng vào khoa học xã hội, phương pháp khoa học chứa đựng một loạt chủ đề như hướng tiếp cận vấn đề nghiên cứu, các công cụ và kỹ thuật thu thập dữ liệu định tính và định lượng, phân tích thống kê, thực nghiệm, khảo sát thực địa, nghiên cứu trường hợp, và v.v... Phần lớn nội dung trong cuốn sách này trình bày về các chủ đề này ở cấp độ thực nghiệm, ví dụ như làm thế nào để quan sát, phân tích và giải thích những dữ liệu thực nghiệm. Chỉ một phần nhỏ trong số các chủ đề này liên quan trực tiếp đến cấp độ lý thuyết, là cấp độ phức tạp hơn rất nhiều trong nghiên cứu khoa học.

Các loại nghiên cứu khoa học

Tùy thuộc vào mục đích nghiên cứu, các công trình nghiên cứu khoa học có thể được phân thành ba loại: nghiên cứu thăm dò, mô tả và giải thích.

Nghiên cứu thăm dò (exploratory research) thường được tiến hành trong các lĩnh vực mới với mục tiêu nghiên cứu là: (1) xác định phạm vi hoặc mức độ của một hiện tượng, một vấn đề hoặc trạng thái nào đó, (2) để hình thành một ý tưởng ban đầu (hoặc "linh cảm") về hiện tượng, hoặc (3) để kiểm tra tính khả thi của việc thực hiện một nghiên cứu quy mô lớn hơn về hiện tượng đó. Ví dụ, nếu công dân của một quốc gia nói chung không hài lòng với chính sách của chính phủ trong giải quyết tình trạng suy thoái kinh tế, thì nghiên cứu thăm dò có thể được sử dụng để đo lường mức độ không hài lòng của công dân, đánh giá biểu hiện sự không hài lòng đó, chẳng hạn như tần suất các cuộc biểu tình công cộng; nguyên nhân không hài lòng chẳng hạn như do chính sách không hiệu quả của chính phủ trong đối phó với lạm phát, lãi suất, thất nghiệp hoặc do thuế má tăng cao. Nghiên cứu thăm dò có thể được tiến hành bằng cách phân tích các số liệu báo cáo công khai của các tổ chức thống kê kinh tế, như các chỉ số kinh tế, tổng sản phẩm quốc nội (GDP), chỉ số thất nghiệp, chỉ số giá tiêu dùng; phân tích ý kiến của các chuyên gia, các

nhà kinh tế hàng đầu, các quan chức chính phủ hoặc phân tích số liệu có được từ các nghiên cứu về những vấn đề tương tự trong lịch sử. Nghiên cứu thăm dò có thể không đưa đến sự hiểu biết chính xác về vấn đề nghiên cứu, nhưng có giá trị trong việc xác định tính chất và mức độ của vấn đề và tạo tiền đề cho nghiên cứu chuyên sâu hơn.

Nghiên cứu mô tả (Descriptive research) là tiến hành quan sát kỹ lưỡng và đưa ra tài liệu chi tiết về một hiện tượng quan tâm. Những quan sát này phải được thực hiện theo phương pháp khoa học (tức là phải có tính lặp lại, tính chính xác,...) và do đó đáng tin cậy hơn những quan sát thông thường. Ví dụ về các nghiên cứu mô tả là bảng số liệu thống kê nhân khẩu học của Cục Điều tra Dân số hoặc bảng thống kê việc làm của Bộ Lao động Hoa Kỳ. Ở đây, người ta sử dụng cùng công cụ hoặc công cụ tương tự để tính toán mức độ gia tăng dân số của các nhóm nhân khẩu khác nhau hoặc tính toán tỉ lệ việc làm theo từng lĩnh vực thông qua các cuộc tổng điều tra dân số hoặc nhiều cuộc khảo sát về lao động, việc làm khác nhau. Nếu thay đổi công cụ đo lường, kết quả thu được có thể sẽ chỉ ra hoặc không chỉ ra được sự thay đổi khi so sánh giữa thời điểm trước và sau về xu hướng diễn biến về dân số hoặc việc làm tại Hoa Kỳ. Một ví dụ khác, nghiên cứu mô tả là các báo cáo xã hội học định kỳ về các hoạt động theo nhóm của trẻ vị thành niên ở khu vực đô thị; về sự duy trì và phát triển các hoạt động tôn giáo, văn hóa trong các cộng đồng dân tộc nào đó; về vai trò của công nghệ như mạng xã hội Twitter và tin nhắn mạng đối với sự lan truyền các phong trào dân chủ tại các quốc gia Trung Đông.

Nghiên cứu giải thích (Explanatory research) tìm cách giải thích các hiện tượng, các vấn đề nảy sinh hoặc trạng thái của sự vật. Trong khi nghiên cứu mô tả xem xét khía cạnh cái gì, ở đâu và khi nào của một hiện tượng thì nghiên cứu giải thích tìm kiếm các câu trả lời cho câu hỏi tại sao và làm thế nào. Nó cố gắng để "kết nối các điểm" trong nghiên cứu bằng cách xác định các yếu tố nguyên nhân và kết quả của hiện tượng. Ví dụ như tìm ra nguyên nhân của tội phạm vị thành niên hoặc nguyên nhân của vấn đề bạo lực của các băng đảng, làm cơ sở xây dựng các chiến lược giải quyết những ung nhọt xã hội này. Hầu hết các nghiên cứu hàn lâm hoặc nghiên cứu tiến sĩ thuộc thể loại giải thích, mặc dù số lượng nghiên cứu thăm dò hoặc mô tả cũng có thể cần thiết trong giai đoạn đầu của nghiên cứu hàn lâm.

Tìm lời giải thích cho các sự kiện được quan sát đòi hỏi phải có kỹ năng lý thuyết và giải thích cùng với khả năng trực giác, sự hiểu biết và kinh nghiệm cá nhân. Những người nghiên cứu có đủ các phẩm chất này là các nhà khoa học có uy tín cao trong chuyên ngành của họ.

Lịch sử tư tưởng khoa học

Trước khi kết thúc chương này, cần lược lại lịch sử phát triển của khoa học qua các thời kỳ và nhận biết những nhà khoa học lỗi lạc trong tiến trình này. Điều này mang lại khá nhiều thú vị. Mặc dù dữ liệu về tiến bộ khoa học được ghi nhận trong suốt nhiều thế kỷ, nhưng thuật ngữ “khoa học”, “nhà khoa học” và “phương pháp khoa học” mới chỉ được đưa ra trong thế kỷ XIX. Trước thời kỳ này, khoa học được xem như là một phần của triết học, tồn tại cùng với các ngành khác của triết học như logic học, siêu hình học, đạo đức học và mỹ học, cho dù ranh giới giữa các ngành này còn mờ nhạt.

Trong thời kỳ bình minh của loài người, tri thức thường được biết đến dưới dạng thuật ngữ giáo huấn thần học dựa trên đức tin. Trong suốt thế kỷ thứ III trước Công nguyên, các nhà triết học Hy Lạp như Plato, Aristotle, Socrates đã phủ nhận điều đó khi các ông cho rằng bản chất cơ bản của hiện hữu và thế giới có thể được hiểu chính xác hơn thông qua quá trình suy luận logic có hệ thống, được gọi là **chủ nghĩa duy lý** (Rationalism). Đặc biệt, công trình nghiên cứu Siêu hình học cổ điển (theo nghĩa đen có nghĩa là "vượt qua [tồn tại] vật lý") của Aristotle (322-384 tr.CN) phân tách thần học (nghiên cứu các vị thần) khỏi bản thể học (nghiên cứu bản thể và tồn tại) và khoa học phổ quát (các nghiên cứu nguyên tắc sơ khởi làm nền tảng của logic). Chủ nghĩa duy lý (không được nhầm lẫn với sự “hợp lý”) xem lý tính là nguồn gốc của tri thức hoặc biện giải và cho rằng tiêu chuẩn của chân lý (truth) không phải là cảm giác mà là lý trí và suy luận diễn dịch - thường bắt nguồn từ một tập hợp các nguyên tắc đầu tiên hoặc tiên đề (chẳng hạn như “Định luật không mâu thuẫn” của Aristotle).

Sự thay đổi quan trọng tiếp theo trong tư tưởng khoa học diễn ra trong suốt thế kỷ XVI, khi nhà triết học người Anh Francis Bacon (1561-1626) đề xuất rằng tri thức chỉ có thể hình thành từ những quan sát trong thế giới thực. Dựa trên tiên đề này, Bacon nhấn mạnh tri thức chỉ có được thông qua hoạt động thực nghiệm (mà không phải là hoạt động tư biện) và phát triển **chủ nghĩa thực nghiệm** (empiricism) thành một nhánh triết học lớn. Các tác phẩm của Bacon dẫn đến (1) sự phổ biến của *phương pháp quy nạp* trong nghiên cứu khoa học, (2) sự phát triển của “phương pháp khoa học” (ban đầu gọi là “phương pháp Bacon”) bao gồm các quan sát có hệ thống, đo lường, thí nghiệm và (3) gieo mầm cho chủ nghĩa vô thần (từ chối các giáo huấn thần học bởi chúng “không thể quan sát được”).

Chủ nghĩa thực nghiệm tiếp tục xung đột với chủ nghĩa duy lý trong suốt thời Trung cổ trong quá trình các nhà triết học tìm cách hữu hiệu nhất để có được tri thức khoa học có

giá trị. Triết gia Pháp Rene Descartes đứng về phía các nhà duy lý, trong khi các nhà triết học người Anh John Locke và David Hume đứng về phía các nhà thực nghiệm chủ nghĩa. Các nhà khoa học khác như Galileo Galilei và Issac Newton đã cố gắng kết hợp hai trường phái này thành triết học học tự nhiên – natural philosophy (triết lý về thế giới tự nhiên), tập trung đặc biệt vào nhiệm vụ tìm hiểu thế giới tự nhiên và vũ trụ - đây được coi là tiền thân của khoa học tự nhiên. Có lẽ Galileo (1564-1642) là người đầu tiên nêu rõ rằng các quy luật của tự nhiên là toán học và là người đã đóng góp vào lĩnh vực thiên văn học thông qua một sự kết hợp sáng tạo giữa các thí nghiệm và toán học.

Vào thế kỷ XVIII, triết gia Đức Immanuel Kant tìm cách giải quyết tranh chấp giữa chủ nghĩa thực nghiệm và chủ nghĩa duy lý. Trong cuốn sách *Phê phán lý tính thuần túy*, ông lập luận rằng kinh nghiệm là hoàn toàn chủ quan và việc xử lý chúng bằng cách sử dụng lý trí thuần túy mà không tìm hiểu sâu bản chất chủ quan của kinh nghiệm sẽ dẫn ảo tưởng lý thuyết. Ý tưởng của Kant đã dẫn đến sự phát triển của **chủ nghĩa duy tâm Đức** (German idealism), truyền cảm hứng phát triển các phương pháp nghiên cứu diễn giải như hiện tượng học, thông diễn học (chú giải văn bản cổ) và lý thuyết xã hội phê phán.

Cũng vào khoảng thời gian đó, nhà triết học Pháp Auguste Comte (1798-1857), người sáng lập ngành xã hội học, đã cố gắng dung hợp chủ nghĩa duy lý và chủ nghĩa thực nghiệm trong một học thuyết mới gọi là **chủ nghĩa thực chứng** (positivism). Ông đề xuất quan điểm rằng lý thuyết và quan sát phụ thuộc nhau theo chu trình hình tròn. Lý thuyết được tạo ra qua quá trình suy luận lý tính, nhưng chúng chỉ được xem là xác thực nếu được kiểm nghiệm thông qua quan sát thực tiễn. Sự nhấn mạnh sự cần thiết của việc kiểm nghiệm lý thuyết qua thực tiễn đã đưa đến sự tách khoa học hiện đại ra khỏi triết học và siêu hình học, phát triển các “phương pháp khoa học” như là phương tiện chủ yếu để kiểm chứng các luận điểm khoa học. Ý tưởng của Comte được Emile Durkheim phát triển thành chủ nghĩa thực chứng trong nghiên cứu xã hội học (thực chứng như là nền tảng của nghiên cứu xã hội) và Ludwig Wittgenstein phát triển thành *chủ nghĩa thực chứng logic* (logical positivism).

Vào đầu thế kỷ XX, các luận điểm cơ bản của chủ nghĩa thực chứng bị các nhà xã hội học diễn giải (các nhà xã hội học phản thực chứng - antipositivists) thuộc trường phái chủ nghĩa duy tâm Đức phản đối. Chủ nghĩa thực chứng thường bị đánh đồng với các phương pháp nghiên cứu định lượng như thí nghiệm và khảo sát mà không có bất kỳ ràng buộc triết lý rõ ràng nào. Trong khi đó, **chủ nghĩa phản thực chứng** (antipositivism) sử dụng các phương pháp định tính như các cuộc phỏng vấn phi cấu trúc và quan sát tham dự.

Ngay cả những người theo chủ nghĩa thực chứng, chẳng hạn nhà xã hội học người Mỹ Paul Lazarsfeld - người đi tiên phong trong nghiên cứu khảo sát quy mô lớn và sử dụng kỹ thuật thống kê để phân tích dữ liệu điều tra, cũng thừa nhận những vấn đề tiềm ẩn liên quan đến định kiến chủ quan của nhà nghiên cứu và những hạn chế ở mức độ vĩ mô của các nghiên cứu thực chứng. Chính vì vậy, các nhà nghiên cứu theo chủ nghĩa phản thực chứng nhấn mạnh các hành vi cá nhân phải được nghiên cứu bằng phương pháp diễn giải, dựa trên sự hiểu được biết về ý nghĩa và mục đích cá nhân chi phối hành vi của họ. Quan điểm này truyền cảm hứng cho công trình nghiên cứu của Georg Simmel về thuyết tương tác biểu tượng (symbolic interactionism), công trình nghiên cứu của Max Weber về các loại hình tư tưởng (ideal types) và công trình nghiên cứu của Edmund Husserl về hiện tượng luận (phenomenology).

Đến giữa và cuối thế kỷ XX, cả hai trường phái thực chứng và phản thực chứng trở thành chủ đề bị chỉ trích và phê phán. Nhà triết học Karl Popper người Anh cho rằng kiến thức nhân loại không dựa trên nền tảng cố định bất biến, mà phải dựa trên một tập hợp các giả thuyết phỏng đoán. Những giả thuyết này có thể không bao giờ được chứng thực một cách chắc chắn và tuyệt đối, mà ngược lại chúng chỉ có thể bị bác bỏ mà thôi. Bằng chứng thực nghiệm là cơ sở để bác bỏ những giả thuyết phỏng đoán hay những “lý thuyết” này. Quan điểm siêu lý thuyết này, được gọi là **thuyết hậu thực chứng** (postpositivism) hay hậu thực nghiệm (postempiricism), đã sửa đổi thuyết thực chứng khi cho rằng chúng ta chỉ có thể bác bỏ một quan điểm sai lầm chứ không thể kiểm chứng được tính chân thực của nó. Tuy nhiên, thuyết hậu thực chứng vẫn thống nhất với thuyết thực chứng về sự tồn tại của sự thật khách quan và nhấn mạnh vai trò của các phương pháp khoa học trong việc hình thành tri thức khoa học.

Trương tợ như vậy, các nhà phản thực chứng cũng bị chỉ trích vì họ chỉ cố gắng để hiểu về xã hội mà không phê phán và đưa ra những thay đổi để xã hội phát triển theo hướng tốt hơn. Nguồn gốc của tư tưởng này nằm trong bộ sách Tư bản (Das Capital) của hai nhà triết học Đức Karl Marx và Friedrich Engels. Học thuyết của Marx và Engels đã phê phán xã hội tư bản là xã hội bất bình đẳng và thiếu hiệu năng; đề xuất giải quyết sự bất bình đẳng thông qua đấu tranh giai cấp và cách mạng vô sản. Chủ nghĩa Marx đã truyền cảm hứng cho các cuộc cách mạng xã hội ở các quốc gia như Đức, Ý, Nga và Trung Quốc. Nhưng nhìn chung, chủ nghĩa Marx đã thất bại trong việc xây dựng một xã hội bình đẳng như mong muốn. Nghiên cứu phê phán (còn gọi là *Lý thuyết phê phán - Critical Research hay Critical theory*) do Max Horkheimer và Jurgen Habermas đề xuất trong thế kỷ XX.

Trong khi giữ lại những tư tưởng phê phán và cách thức giải quyết bất bình đẳng xã hội tương tự chủ nghĩa Marx, *Lý thuyết phê phán* bổ sung rằng con người có khả năng và cần hành động có ý thức để thay đổi hoàn cảnh kinh tế và địa vị xã hội của mình, mặc dù khả năng thực hiện của họ bị hạn chế bởi các hình thức khác nhau của sự thống trị về chính trị, văn hóa và xã hội. Học thuyết phê phán cố gắng hướng tới phát hiện và phê phán các điều kiện khó khăn, trở ngại của hoàn cảnh thực tại bằng cách phân tích các đối kháng, xung đột và mâu thuẫn trong xã hội đương đại, tìm cách loại bỏ nguồn gốc của sự tha hóa và sự thống trị (tức là giải phóng giai cấp bị áp bức). Còn nhiều triết lý nghiên cứu và phương pháp tiếp cận khác nhau sẽ được đề cập trong các chương tiếp theo của cuốn sách này.

Chương 2

TƯ DUY CỦA NHÀ NGHIÊN CỨU

Muốn tiến hành nghiên cứu được tốt đòi hỏi trước hết phải rèn luyện lại bộ não của bạn để tư duy như một nhà nghiên cứu thực thụ. Muốn vậy phải trừu tượng hóa những quan sát thực tiễn, sử dụng tư duy "liên kết các điểm" để nhận diện các khái niệm, các mô hình ẩn và sau đó tổng hợp những mô hình đó trở thành quy luật và lý thuyết phổ quát để áp dụng trong các bối cảnh khác vượt ra ngoài lĩnh vực nghiên cứu ban đầu. Nghiên cứu là việc liên tục vận động hoán đổi từ *mức độ thực nghiệm* khi tiến hành các quan sát lên *mức độ lý thuyết* khi những quan sát được trừu tượng hóa thành các quy luật và lý thuyết phổ quát. Đây là kỹ năng phải mất nhiều năm để phát triển và cho đến nay, nó không phải là cái được giảng dạy trong các chương trình sau đại học hay có được trong đào tạo thực tập và đó chính là thiếu hụt lớn nhất đối với các nghiên cứu sinh tiến sĩ. Một số khái niệm trừu tượng cần phải được tiếp cận bằng tư duy của nhà nghiên cứu, bao gồm: đơn vị phân tích, phạm trù, giả thuyết, thao tác hóa, lý thuyết, mô hình, phương pháp quy nạp, diễn dịch,... Tất cả sẽ được đề cập trong chương này.

Đơn vị phân tích

Một trong những thuật ngữ đầu tiên xuất hiện trong tất cả nghiên cứu khoa học xã hội là *đơn vị phân tích* (unit of analysis) của một nghiên cứu khoa học. Đơn vị phân tích đề cập đến người, tập thể hoặc vật thể là đối tượng mà nghiên cứu hướng tới. Đơn vị điển hình trong phân tích bao gồm các cá nhân, các nhóm, các tổ chức, các quốc gia, các công nghệ, các đồ vật,... Ví dụ, nếu quan tâm đến nghiên cứu hoạt động mua sắm của người dân, mục đích mua sắm của họ hoặc thái độ của họ với công nghệ mới thì đơn vị phân tích là cá nhân. Nếu nghiên cứu đặc điểm của các băng nhóm tội phạm đường phố hoặc hoạt động làm việc theo nhóm trong các tổ chức thì đơn vị phân tích là nhóm. Nếu mục đích của nghiên cứu là tìm hiểu xem làm thế nào các doanh nghiệp có thể nâng cao lợi nhuận hoặc đưa ra các quyết định điều hành hiệu quả thì đơn vị phân tích là công ty. Trong trường hợp này, mặc dù quyết định được thực hiện bởi các cá nhân trong công ty, nhưng những cá nhân được coi là đại diện cho việc ra quyết định của công ty chứ không phải là quyết định của cá nhân họ. Nếu nghiên cứu hướng vào những khác biệt về các nền văn hóa dân tộc thì đơn vị phân tích lại là một quốc gia. Ngay cả những vật vô tri, vô giác có thể trở thành

những đơn vị phân tích. Ví dụ, một nhà nghiên cứu quan tâm đến việc tìm hiểu giải pháp để làm cho các trang web hấp dẫn hơn đối với người sử dụng, thì đơn vị phân tích là một trang web (mà không phải người sử dụng). Khi nghiên cứu cách thức chuyển giao tri thức giữa hai công ty, thì đơn vị phân tích phân tích trở thành nhị nguyên (sự kết hợp của các công ty giao và nhận tri thức).

Hiểu biết về các đơn vị phân tích đôi khi khá phức tạp. Ví dụ, nếu chúng ta muốn nghiên cứu lý do tại sao một số khu vực nào đó có tỷ lệ tội phạm cao thì đơn vị phân tích của chúng ta là các khu phố mà không phải là các loại tội phạm hay người có hành vi phạm tội. Điều này là do các đối tượng cần khảo sát là các khu phố mà không phải là kẻ phạm tội. Tuy nhiên, nếu chúng ta muốn so sánh các loại tội phạm khác nhau trong các khu vực khác nhau, chẳng hạn như các tội giết người, cướp tài sản, hành hung,... thì đơn vị phân tích lại là các hành vi phạm tội. Nếu muốn nghiên cứu lý do tại sao bọn tội phạm tham gia vào các hoạt động bất hợp pháp, thì đơn vị phân tích là cá nhân (người có hành vi phạm tội). Tương tự như vậy, nếu muốn nghiên cứu tại sao một số chính sách đổi mới lại có nhiều thành công hơn những chính sách khác, thì đơn vị của phân tích là một chính sách đổi mới. Tuy nhiên, nếu chúng ta muốn nghiên cứu bằng cách nào một số tổ chức lại có những đổi mới nhất quán và liên tục hơn so với những tổ chức khác, thì đơn vị phân tích là tổ chức. Do đó, hai câu hỏi liên quan trong cùng một nghiên cứu có thể có hai đơn vị phân tích hoàn toàn khác nhau.

Hiểu biết về các đơn vị phân tích là rất quan trọng, bởi vì nó sẽ quy định loại thông tin, tài liệu nào nào bạn cần thu thập và thu thập từ đâu. Nếu đơn vị phân tích là một trang web, bạn nên thu thập dữ liệu từ các trang web chứ không phải là khảo sát mọi người về cách họ sử dụng các trang web. Nếu đơn vị bạn phân tích là tổ chức thì bạn nên đánh giá các thông số của tổ chức, chẳng hạn như quy mô của tổ chức, doanh thu, hệ thống phân cấp hoặc khả năng phát triển. Dữ liệu này có thể đến từ nhiều nguồn như các hồ sơ tài chính, khảo sát từ các cán bộ điều hành, Tổng giám đốc điều hành (CEO), người được cho là đại diện cho tổ chức của họ (chứ không phải bản thân công ty). Một số thông số như lương của Tổng giám đốc điều hành có vẻ giống như các thông số ở cấp độ cá nhân, nhưng trên thực tế, nó cũng có thể là một thông số ở cấp độ tổ chức, bởi vì mỗi tổ chức chỉ có một lương cho Tổng giám đốc điều hành tại một thời điểm bất kỳ nào đó. Đôi khi, bạn có thể thu thập dữ liệu từ một mức độ thấp hơn, sau đó phân tích và tổng hợp dữ liệu đến một cấp độ cao hơn. Chẳng hạn, để nghiên cứu nhóm làm việc trong các tổ chức, bạn có thể khảo sát thành viên của nhóm trong các tổ chức khác và sau đó lấy điểm số trung bình các cá nhân của

nhóm để tạo ra một biến số nhóm với sự gắn kết và khác biệt. Chúng ta sẽ xem xét khái niệm “biến số” kỹ hơn trong phần tiếp theo.

Khái niệm, phạm trù và biến số

Như đã trình bày ở Chương 1, mặc dù nghiên cứu có thể được phân chia thành nghiên cứu thăm dò, mô tả hay giải thích, nhưng hầu hết các nghiên cứu khoa học đều chứa đựng phần nào đó nội dung giải thích, bởi vì chúng tìm kiếm lời giải thích cho các hiện tượng tự nhiên hoặc xã hội được quan sát. Những giải thích này đặt ra yêu cầu phải phát triển các **khái niệm** (concept) về các thuộc tính khái quát, đặc điểm liên quan tới đối tượng, sự kiện hoặc con người. Trong khi các đối tượng như một người, một công ty hoặc một chiếc xe hơi không phải là các khái niệm, thì các đặc điểm cụ thể hoặc hành vi của chúng, ví dụ như thái độ của một người đối với người nhập cư, năng lực của một công ty để tạo ra sự sáng tạo, trọng lượng của chiếc xe - có thể được xem như là những khái niệm.

Dù nhận thức rõ hay không, chúng ta đã sử dụng các khái niệm khác nhau trong đời sống hàng ngày. Một vài trong số các khái niệm đó đã phát triển qua thời gian, thông qua sự chia sẻ ngôn ngữ. Đôi khi, chúng ta mượn các khái niệm từ những chuyên ngành hoặc ngôn ngữ khác để giải thích một hiện tượng mà mình quan tâm. Ví dụ, ý tưởng về *lực hấp dẫn* được mượn từ vật lý học có thể được sử dụng trong ngành kinh doanh để mô tả tại sao khách hàng “*bị hút*” vào những điểm đến mua sắm yêu thích của họ. Tương tự như vậy, khái niệm *khoảng cách* có thể được sử dụng để giải thích mức độ chia cách xã hội của hai cá nhân ở các địa vị khác nhau. Đôi khi, chính chúng ta sáng tạo ra những khái niệm mới để mô tả một đặc điểm riêng biệt chưa được mô tả trong các nghiên cứu trước đó. Ví dụ, “*căng thẳng công nghệ*” (technostress) là một khái niệm mới liên quan đến trạng thái tâm lý căng thẳng (stress) của một người khi đối mặt với yêu cầu tiếp nhận một công nghệ mới.

Các khái niệm có thể có những mức độ trừu tượng khác nhau. Một vài khái niệm như *trọng lượng* của một người là rõ ràng và khách quan; trong khi các khái niệm khác như *tính cách* của một người lại có thể trừu tượng và khó để hình dung. **Phạm trù** (construct)^(*) là khái niệm trừu tượng được lựa chọn (hoặc được xây dựng) một cách riêng biệt nhằm mục đích diễn đạt một hiện tượng nhất định. Phạm trù có thể là một khái niệm đơn giản như trọng lượng của một người hoặc là sự kết hợp của một hệ thống các khái niệm có liên

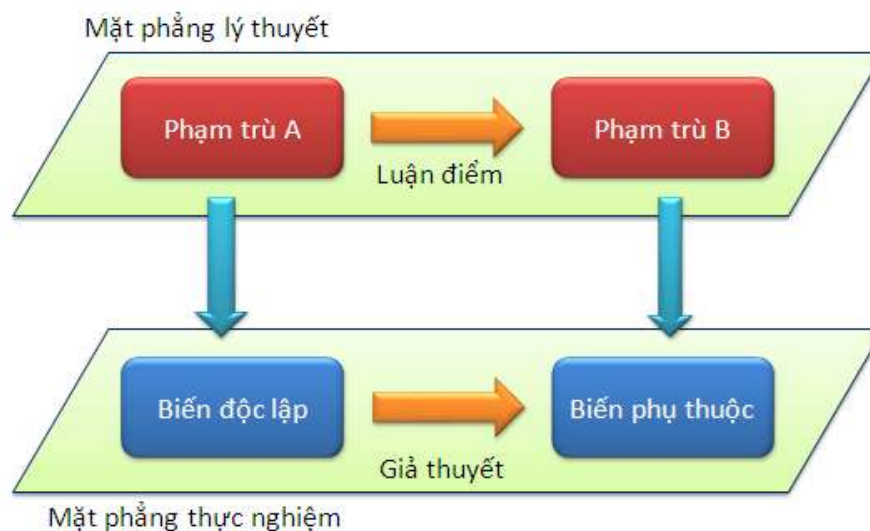
(*) GS, TS Dương Thiệu Tống chú giải “*construct*” là “*khái niệm tạo lập*”, nghĩa là khái niệm “*được nhà nghiên cứu sáng tạo ra một cách có chủ đích và có ý thức, hay được người nghiên cứu chấp nhận và sử dụng cho mục tiêu tìm hiểu khoa học của mình*”. Xem: Phương pháp nghiên cứu khoa học giáo dục và tâm lý, Nxb Khoa học xã hội, 2005, tr.45 (N.D).

quan, ví dụ *kỹ năng giao tiếp* của một người chứa đựng một vài khái niệm cơ sở như vốn từ vựng, cú pháp và khả năng chính tả của một người. Ví dụ về trọng lượng được gọi là **phạm trù đơn** (uni-dimensional construct), trong khi ví dụ về kỹ năng giao tiếp được gọi là **phạm trù phức** (multi-dimensional construct) chứa đựng hàng loạt khái niệm cơ bản khác. Sự khác biệt giữa phạm trù và khái niệm rõ ràng hơn trong các phạm trù đa chiều: khi sự trừu tượng ở mức độ cao hơn được gọi là một phạm trù và sự trừu tượng ở mức độ thấp hơn được gọi là khái niệm. Tuy nhiên, sự phân biệt này có thể không rõ ràng trong trường hợp của các phạm trù đơn.

Phạm trù sử dụng cho nghiên cứu khoa học phải được định nghĩa chính xác và rõ ràng, có như vậy người khác mới có thể hiểu đúng những gì nó bao hàm và những gì không được bao hàm. Ví dụ, phạm trù *thu nhập*, có vẻ như đơn giản nhưng có thể ám chỉ thu nhập hàng tháng hoặc hàng năm, thu nhập trước thuế hay sau thuế, thu nhập cá nhân hoặc gia đình. Như vậy, dưới góc độ nghiên cứu khoa học thì phạm trù này không chính xác và cũng không rõ ràng. Có hai loại định nghĩa có thể sử dụng để làm rõ phạm trù: định nghĩa theo từ điển (dictionary definition) và định nghĩa theo thao tác (operational definition). **Định nghĩa theo từ điển** quen thuộc hơn, trong đó phạm trù thường được định nghĩa thông qua hệ thống các từ đồng nghĩa. Ví dụ, *thái độ* có thể được định nghĩa như một thiên hướng, một cảm giác hoặc một cảm xúc; và đến lượt cảm xúc lại có thể được định nghĩa như một thái độ. Những định nghĩa kiểu vòng tròn như vậy không thật sự hữu ích trong nghiên cứu khoa học nếu muốn chi tiết hóa ý nghĩa và nội dung của phạm trù đó. Nghiên cứu khoa học đòi hỏi **định nghĩa theo thao tác** - định nghĩa các phạm trù bằng cách đánh giá chúng qua thực tiễn. Ví dụ, định nghĩa theo thao tác về một phạm trù như *nhiệt độ*, thì phải cụ thể hóa liệu rằng chúng ta có đo được nhiệt độ bằng độ C, độ F hay độ K hay không. Một phạm trù như *thu nhập* nên được định nghĩa dưới góc độ liệu rằng chúng ta có quan tâm tới thu nhập hàng tháng hoặc hàng năm, trước thuế thu nhập sau thuế, thu nhập cá nhân hoặc gia đình. Có thể hình dung những phạm trù như học tập, nhân cách, trí thông minh - cũng rất khó để định nghĩa thao tác.

Một thuật ngữ thường kết hợp và đôi khi được hoán đổi với phạm trù được gọi là một **biến số** (variable). Xét về từ gốc, biến số là số lượng có thể thay đổi (ví dụ, từ thấp đến cao, từ tiêu cực sang tích cực, v.v...), ngược với hằng số không thay đổi (bất biến). Tuy nhiên, trong nghiên cứu khoa học, biến số là một đại lượng có thể đo được của một phạm trù trừu tượng. Trong các thực thể trừu tượng, các phạm trù không thể đo lường trực tiếp được. Do đó, chúng ta phải tìm kiếm các đơn vị đo lường thay thế được gọi là các biến số.

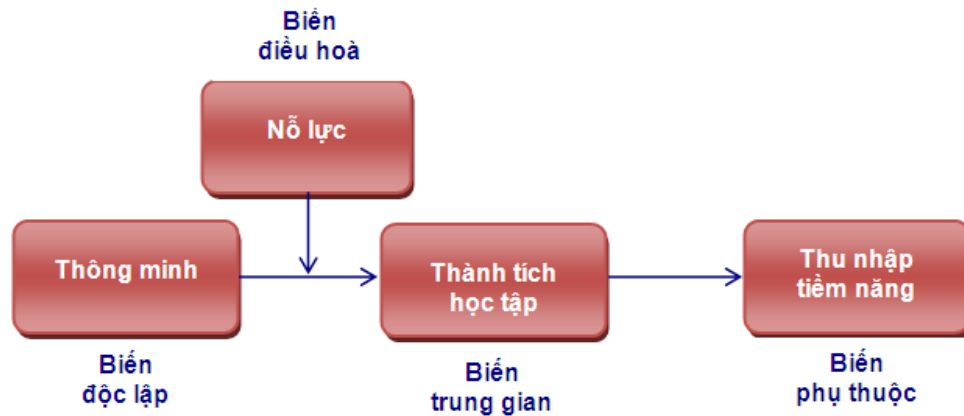
Ví dụ, *trí thông minh* của một người thường được đo lường bằng chỉ số IQ (Intelligence Quotient). Đó là một chỉ số được tạo ra khi phân tích và kiểm tra mẫu dùng cho con người. Trong trường hợp này *trí thông minh* là một phạm trù và chỉ số IQ là một biến số để đo lường phạm trù trí thông minh. Liệu rằng chỉ số IQ có phản ánh đúng sự thông minh của một người như mọi người vẫn nghĩ và tin tưởng? Tùy thuộc vào mức độ chính xác khi đánh giá trí thông minh, chỉ số IQ có thể là một công cụ tốt hoặc không tốt để đánh giá phạm trù trí thông minh. Như trong Hình 2.1, nghiên cứu khoa học phát triển trên 2 mặt phẳng: mặt phẳng lý thuyết và mặt phẳng thực tiễn. Các phạm trù được nhận thức trên mặt phẳng lý thuyết (trừu tượng), trong khi đó các biến số được thao tác và đo lường trên mặt phẳng thực tiễn (quan sát). Tư duy giống như một nhà nghiên cứu là nói tới khả năng tương tác qua lại của tư duy giữa hai mặt phẳng này.



Hình 2.1. Mặt phẳng lý thuyết và mặt phẳng thực nghiệm của nghiên cứu

Tùy thuộc vào mục đích sử dụng mà các biến có thể được phân loại là các biến số độc lập, phụ thuộc, điều hòa, trung gian hoặc kiểm soát. Các biến số được sử dụng để giải thích các biến số khác được gọi là *biến độc lập* (independent variable), những biến được giải thích bởi các biến số khác là *biến phụ thuộc* (dependent variable). Những biến được giải thích bởi các biến độc lập trong khi chúng cũng giải thích các biến phụ thuộc được gọi là *biến trung gian* (mediating variable). Những biến có ảnh hưởng đến mối quan hệ giữa các biến độc lập và biến phụ thuộc được gọi là *biến điều hòa* (moderating variable). Ví dụ, nếu cho rằng trong sinh viên trí thông minh cao hơn sẽ tạo ra kết quả học tập tốt hơn, vậy thì trí thông minh là biến độc lập và kết quả học tập là biến phụ thuộc. Có thể có các biến

số khác không liên quan và không thích hợp để giải thích một biến phụ thuộc, nhưng có thể có một số tác động đến biến phụ thuộc. Các biến này phải được kiểm soát trong nghiên cứu khoa học và do đó được gọi là các *biến kiểm soát* (control variable).



Hình 2.2. Mạng tương tác của các phạm trù

Để hiểu được sự khác nhau giữa các loại biến, hãy xem xét ví dụ thể hiện trong Hình 2.2. Nếu chúng ta tin rằng, trí thông minh ảnh hưởng đến thành tích học tập thì việc đo lường trí thông minh chẳng hạn như chỉ số IQ là biến độc lập. Trong khi đó thước đo thành tích học tập như điểm tổng kết là biến phụ thuộc. Nếu chúng ta tin rằng tác động của trí thông minh đối với thành tích học tập cũng phụ thuộc vào những nỗ lực mà sinh viên đầu tư vào quá trình học tập (ví dụ giữa hai sinh viên thông minh như nhau, sinh viên nào nỗ lực nhiều hơn có thể đạt được thành tích học tập cao hơn), vậy thì nỗ lực trở thành biến điều hòa. Một cách ngẫu nhiên, người ta cũng có thể xem nỗ lực là một biến độc lập và thông minh là một biến điều hòa. Nếu thành tích học tập được xem như là bước trung gian để có thu nhập tiềm năng cao hơn, như vậy thu nhập tiềm năng trở thành biến phụ thuộc vào biến độc lập là thành tích học tập. Do đó, thành tích học tập trở thành các biến trung gian trong mối quan hệ giữa trí thông minh và thu nhập tiềm năng. Do đó, biến số được xem là độc lập, phụ thuộc, kiểm soát hay trung gian dựa vào mối quan hệ giữa chúng. Mạng lưới khái quát của các mối quan hệ giữa một tập hợp của các phạm trù có liên quan được gọi là *mạng tương tác* (nomological network - xem Hình 2.2). Tư duy của nhà nghiên cứu không chỉ đòi hỏi khả năng hình thành các phạm trù từ những quan sát thực tiễn, mà còn yêu cầu hình dung trong tư duy một mạng tương tác để xâu chuỗi và liên kết các phạm trù trừu tượng.

Luận điểm và giả thuyết

Hình 2.2 cho thấy cách thức liên kết giữa các phạm trù lý thuyết như trí thông minh, nỗ lực, thành tích học tập và thu nhập tiềm năng trong một mạng tương tác. Mỗi mối quan hệ này được gọi là một luận điểm. Việc tìm kiếm những giải thích cho một hiện tượng hoặc một hành vi nhất định sẽ là không đầy đủ nếu chỉ xác định các khái niệm và phạm trù cơ bản liên quan đến hiện tượng hay hành vi đó. Chúng ta cũng phải xác định và hình thành các mô hình phản ánh mối quan hệ giữa các phạm trù này. Mô hình các mối quan hệ như vậy được gọi là luận điểm. **Luận điểm** (proposition) là một quan hệ thăm dò và phỏng đoán giữa các phạm trù được trình bày dưới dạng mệnh đề. Một ví dụ về luận điểm là: “*Sự cải thiện trí thông minh của học sinh tạo ra sự cải thiện thành tích học tập của họ*”. Mệnh đề này không bắt buộc phải đúng (có thể đúng, có thể sai), nhưng phải là mệnh đề có thể kiểm chứng được bằng dữ liệu thực nghiệm; và sau khi có kết quả kiểm chứng, chúng ta có thể kết luận nó đúng hay sai. Luận điểm thường được xây dựng dựa trên suy luận logic (diễn dịch) hay thông qua quan sát thực nghiệm (quy nạp).

Do luận điểm là sự kết hợp giữa các phạm trù trừu tượng nên chúng không thể được kiểm chứng trực tiếp. Thay vào đó, chúng được kiểm chứng gián tiếp bằng cách xem xét các mối quan hệ giữa các đơn vị đo lường (các biến số) tương ứng với các phạm trù đó. Sự hình thành các luận điểm bằng thực nghiệm đề cập đến mối quan hệ giữa các biến số, được gọi là các **giả thuyết** (Hypothesis - xem Hình 2.1). Bởi vì chỉ số IQ và điểm tổng kết học tập là các công cụ tương ứng để đánh giá trên thực tế trí thông minh và thành tích học tập, luận điểm đã nêu ở trên có thể được phát biểu cụ thể dưới hình thức là một giả thuyết “*Sự cải thiện điểm số IQ của sinh viên tạo ra sự cải thiện điểm tổng kết học tập của họ*”. Luận điểm được cụ thể hóa trên mặt phẳng lý thuyết, trong khi đó giả thuyết được cụ thể hóa trên mặt phẳng thực tiễn. Vì vậy, các giả thuyết này hoàn toàn có thể kiểm chứng được trong thực tiễn bằng việc sử dụng các dữ liệu đã thu thập và giả thuyết này có thể bị bác bỏ nếu không được minh chứng bởi các quan sát thực nghiệm. Tất nhiên, mục đích của việc kiểm định các giả thuyết là để suy ra luận điểm tương ứng có chính xác hay không.

Giả thuyết có thể mạnh hoặc yếu. “*Chỉ số IQ của sinh viên có liên quan tới thành tích học tập của họ*” là một ví dụ về một giả thiết yếu, bởi vì nó không chỉ rõ được cả định hướng của giả thuyết (ví như liệu rằng mối quan hệ này là tích cực hay tiêu cực) và cả quan hệ nhân - quả của nó (ví như trí thông minh mang đến thành tích học tập hay thành tích học tập gây ra trí thông minh). Một giả thuyết mạnh hơn như là “*Chỉ số IQ của sinh viên có quan hệ tích cực với thành tích học tập của họ*”. Giả thuyết này chỉ ra định hướng

nhưng chưa nêu ra được quan hệ nhân quả. Một giả thuyết tốt hơn nữa sẽ là “*Chỉ số IQ của sinh viên có những ảnh hưởng tích cực đến thành tích học tập của họ*”. Ở đây giả thuyết đã chỉ ra cả định hướng cũng như quan hệ nhân quả (trí thông minh tạo ra thành tích học tập và không ngược lại). Những kí hiệu trong Hình 2.2 chỉ ra định hướng và các giả thuyết tương ứng.

Cần phải chú ý rằng, các giả thuyết khoa học nên xác định rõ những biến số độc lập và phụ thuộc. Trong giả thuyết “*Chỉ số IQ của sinh viên có những ảnh hưởng tích cực đến thành tích học tập của họ*”, thì trí thông minh là biến số độc lập (nguyên nhân) và thành tích học tập là biến số phụ thuộc (kết quả). Hơn nữa, cũng cần thấy rằng, giả thuyết này có thể đúng (trí thông minh cao hơn sẽ dẫn đến thành tích học tập tốt hơn), nhưng cũng có thể sai (trí thông minh cao hơn không có ảnh hưởng hoặc không dẫn đến kết quả học tập tốt hơn). Phần sau của cuốn sách này, chúng ta sẽ kiểm tra cách kiểm định thực nghiệm các mối quan hệ nguyên nhân - kết quả. Những phát biểu như “*sinh viên nhìn chung là thông minh*” hoặc “*tất cả sinh viên có thể đạt được những thành công trong học tập*” không phải là những giả thuyết khoa học, bởi lẽ chúng không chỉ rõ các biến số độc lập và phụ thuộc, không chỉ rõ quan hệ định hướng để đánh giá nó đúng hay sai.

Lý thuyết và mô hình

Lý thuyết (theory) là tập hợp các phạm trù và luận điểm có quan hệ tương hỗ nhằm phán đoán, giải thích một hiện tượng hoặc một hành vi cần quan tâm trong phạm vi một số điều kiện và giả thiết nhất định. Về bản chất, lý thuyết là một tập hợp có hệ thống những luận điểm. Trong khi luận điểm liên kết hai hoặc ba phạm trù thì lý thuyết tương ứng với một *hệ thống* các phạm trù và luận điểm phức hợp. Chính vì vậy, lý thuyết có thể phức tạp và trừu tượng hơn rất nhiều so với luận điểm và giả thuyết.

Chú ý rằng, một số người không được trang bị kiến thức về nghiên cứu khoa học thường coi lý thuyết như là một *sự suy đoán* (speculation) hay là cái đối lập với *thực tế* (fact). Ví dụ, người ta thường nói rằng giáo viên cần phải giảm lý thuyết và tăng thực hành hay thực tế trong các bài giảng của họ trên lớp. Tuy nhiên, thực tế (fact) hay thực tiễn (practice) không đối lập với lý thuyết. Trong tư duy khoa học, chúng là những thành phần cơ bản, cần thiết để kiểm tra giá trị của một lý thuyết. Một lý thuyết khoa học tốt được hỗ trợ mạnh mẽ bởi các bằng chứng thực tế được quan sát; và một lý thuyết khoa học tốt cũng cần có giá trị thực tiễn cao. Và như vậy, một lý thuyết yếu thường thiếu vắng những yếu tố này. Nhà nghiên cứu về tổ chức nổi tiếng Kurt Lewin (1945) đã nói rằng “Lý thuyết mà

không có thực tiễn là lý thuyết vô ích; thực tiễn không có lý thuyết là thực tiễn mù quáng^(*). Vì vậy, lý thuyết và thực tế (hay thực tiễn) là nền tảng cho nghiên cứu khoa học.

Các lý thuyết cung cấp giải thích về các hiện tượng tự nhiên và xã hội. Như đã nhấn mạnh ở Chương 1, các giải thích này có thể thuyết phục hoặc không thuyết phục, vì vậy, có những lý thuyết tốt và kém. Chương 3 sẽ mô tả một số tiêu chí để đánh giá một lý thuyết. Tuy nhiên, có một điều quan trọng đối với các nhà nghiên cứu đó là lý thuyết không phải là “chân lý” (truth), không có lý thuyết nào là thần thánh, bất khả xâm phạm và các lý thuyết không nên được chấp nhận chỉ vì chúng được đề xuất bởi một ai đó. Trong tiến trình khoa học, các lý thuyết lỗi thời cuối cùng sẽ bị thay thế bởi các lý thuyết tân tiến với khả năng giải thích thuyết phục hơn. Thách thức cơ bản đối với các nhà nghiên cứu là làm sao để xây dựng được những lý thuyết tốt hơn, hoàn thiện hơn và chúng có thể giải thích các hiện tượng đang quan tâm tốt hơn các lý thuyết trước đó.

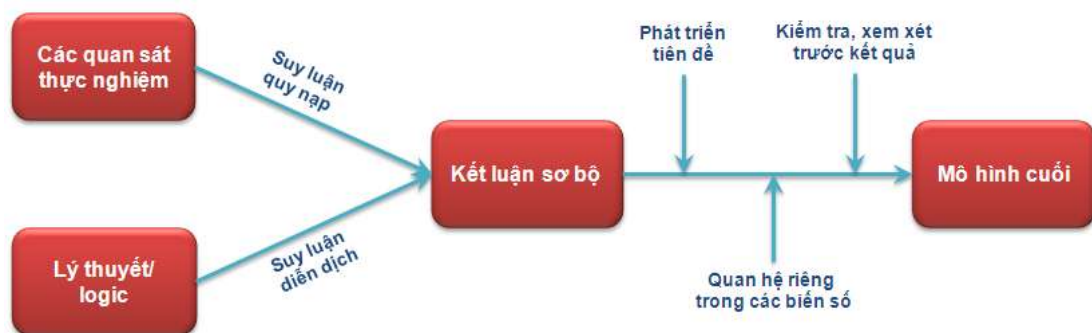
Một thuật ngữ thường được sử dụng liên với lý thuyết là mô hình. **Mô hình** (model) là sự hiện diện tất cả hoặc một phần của một hệ thống được xây dựng để nghiên cứu chính hệ thống đó (ví dụ nghiên cứu cách thức hoạt động hay các tâm điểm của hệ thống đó). Trong khi lý thuyết cố gắng giải thích một hiện tượng, thì mô hình lại cố gắng đại diện (mô tả) cho một hiện tượng. Dựa trên nguồn đầu vào, người nghiên cứu hình thành nên các mô hình để thực hiện những nhiệm vụ quan trọng trong nghiên cứu. Ví dụ, những nhà quản lý thị trường trong doanh nghiệp có thể sử dụng các mô hình để quyết định lượng tiền sẽ sử dụng vào việc quảng cáo cho những dòng sản phẩm khác nhau dựa trên các thông số về chi tiêu cho quảng cáo năm trước, doanh số, sự phát triển của thị trường và các sản phẩm đang cạnh tranh với họ. Tương tự như vậy, các nhà dự báo thời tiết có thể sử dụng các mô hình để dự báo các loại thời tiết trong tương lai dựa trên các thông số như tốc độ và hướng gió, nhiệt độ, độ ẩm. Trong khi các mô hình này đang được sử dụng, không chỉ là giải thích chi phí quảng cáo hay dự báo thời tiết; mô hình có nhiều loại, như mô hình toán học, mô hình mạng lưới, mô hình đường dẫn. Mô hình có thể đảm nhận chức năng mô tả, dự báo hoặc quy chuẩn. Mô hình mô tả thường dùng đại diện các hệ thống phức tạp nhằm biểu thị các biến số và quan hệ trong hệ thống đó. Mô hình chi phí quảng cáo có thể coi là mô hình mô tả. Mô hình dự báo (ví dụ mô hình hồi quy) cho phép dự báo các sự kiện trong tương lai. Mô hình dự báo thời tiết thuộc mô hình dự báo. Mô hình quy chuẩn được dùng để hướng

(*) Luận điểm: “*Practice without theory is blind. Theory without practice is sterile*” (Karl Marx) ở Việt Nam thường được dịch là: Thực tiễn không có lý luận hướng dẫn là thực tiễn mù quáng; lý luận mà không liên hệ với thực tiễn là lý luận suông (N.D).

dẫn hoạt động của chúng ta theo các chuẩn mực phổ dụng trong thực tiễn. Mô hình có thể là tĩnh nếu biểu thị trạng thái của hệ thống ở một thời điểm hoặc là mô hình động nếu biểu thị sự phát triển của hệ thống suốt thời gian dài.

Tiến trình phát triển của mô hình và lý thuyết có thể bao gồm suy luận quy nạp và diễn dịch. Trở lại Chương 1, **suy luận diễn dịch** (deduction) là một quá trình tìm ra kết luận về một hiện tượng hoặc hành vi dựa trên nền tảng lý thuyết hoặc suy luận logic từ một tập hợp các tiên đề. Ví dụ, nếu một ngân hàng thực thi chính sách chặt chẽ về đạo đức của nhân viên (tiên đề 1), Jamie là một nhân viên ở ngân hàng đó (tiên đề 2) thì Jamie phải tuân theo những nguyên tắc đạo đức đó (kết luận). Trong suy luận diễn dịch, các kết luận chắc chắn sẽ chính xác nếu như các tiên đề và các suy luận đều chính xác.

Ngược lại, **suy luận quy nạp** (induction) là quá trình đưa ra kết luận dựa trên các sự kiện thực tế và chứng cứ đã quan sát, thu thập được. Ví dụ, nếu một công ty chi tiêu rất nhiều tiền cho các đợt khuyến mại (Quan sát 1), nhưng doanh số lại không tăng (Quan sát 2), vậy thì có khả năng các đợt khuyến mại không được tổ chức tốt (Kết luận). Tuy nhiên, có thể có các giải thích khác cho sự giảm doanh số, ví dụ do suy thoái kinh tế, sự xuất hiện của các sản phẩm cạnh tranh hoặc chuỗi nguồn cung có vấn đề. Vậy thì, những kết luận quy nạp chỉ là một giả thuyết và có thể bị bác bỏ. Các kết luận diễn dịch nhìn chung có giá trị vững chắc hơn các kết luận quy nạp, nhưng kết luận diễn dịch dựa trên một tiên đề sai thì cũng sai.



Hình 2.3. Quá trình xây dựng mô hình

Như trong Hình 2.3, suy luận quy nạp và diễn dịch có mối quan hệ mật thiết với việc xây dựng lý thuyết và mô hình. Sự quy nạp xuất hiện khi chúng ta quan sát một thực tế và hỏi “Tại sao điều này xảy ra?”. Để trả lời, chúng ta phát triển một hoặc vài giải thích thăm dò (các giả thuyết). Sau đó chúng ta sử dụng các phương pháp diễn dịch, dựa trên các tiên

đề hợp lý và có cơ sở là những hiểu biết về hiện tượng đang nghiên cứu để thu nhỏ các giải thích thăm dò thành một giải thích đáng tin nhất. Để tìm kiếm một sự mở rộng, phát triển hay thay đổi so với lý thuyết hoặc mô hình hiện có, nhà nghiên cứu phải có khả năng tương tác những suy luận quy nạp và diễn dịch. Đây chính là bản chất của nghiên cứu khoa học.

Chương 3

TIẾN TRÌNH NGHIÊN CỨU

Ở Chương 1 chúng ta đã hiểu rằng nghiên cứu khoa học là một quá trình tìm kiếm những tri thức khoa học bằng việc sử dụng các phương pháp khoa học. Tuy nhiên, nghiên cứu đó được thực hiện như thế nào? Chương này sẽ phân tích kỹ lưỡng tiến trình của một nghiên cứu khoa học, các giả thiết và kết quả của tiến trình nghiên cứu đó.

Các mô thức trong nghiên cứu xã hội

Việc thiết kế và thực hiện nghiên cứu bị chi phối bởi các mô hình trong tư duy (mental models) hoặc những khung lý thuyết hay còn gọi là hệ quy chiếu (frames of references) mà chúng ta sử dụng để tổ chức các luận điểm và quan sát. Những mô hình tư duy và khung lý thuyết (hệ thống niềm tin) này được gọi chung là các *mô thức* (paradigm). Thuật ngữ “mô thức” được Thomas Kuhn (1962) phổ biến rộng rãi trong cuốn sách của ông mang tên “*Cấu trúc của các cuộc cách mạng khoa học*” (*The Structure of Scientific Revolutions*). Tại đây, ông đã nghiên cứu lịch sử của khoa học tự nhiên để tìm ra các khung mẫu hành vi khác nhau có ảnh hưởng ở mức độ khác nhau đến tiến bộ khoa học.

Tương tự như vậy, trong khoa học xã hội, một hiện thực xã hội có thể được nhìn nhận khác nhau bởi những người khác nhau phụ thuộc vào cách tư duy và nhận thức của họ về thực tại đó. Ví dụ, những người theo khuynh hướng bảo thủ và khuynh hướng tự do thường có những quan niệm rất khác biệt về vai trò của chính phủ trong đời sống của người dân. Do vậy, họ giữ những quan điểm rất khác nhau về cách xử lý các vấn đề xã hội. Những người bảo thủ tin rằng cắt giảm thuế là cách tốt nhất để vực dậy một nền kinh tế suy thoái, bởi vì nó làm gia tăng thu nhập sau thuế và tiêu dùng của người dân. Điều này tác động trở lại làm tăng đầu ra cho các ngành kinh doanh cũng như tạo thêm công ăn việc làm. Ngược lại, những người theo chủ nghĩa tự do lại cho rằng, chính phủ nên đầu tư nhiều hơn vào các chương trình tạo việc làm, như các công việc công cộng và các đề án nâng cấp cơ sở hạ tầng. Điều này sẽ tạo ra nhiều việc làm và tăng khả năng chi tiêu của người dân, từ đó tác động tích cực đến nền kinh tế. Tương tự như vậy, xã hội phương Tây hướng trọng tâm nhiều hơn vào các quyền cá nhân, ví dụ như quyền riêng tư, quyền tự do ngôn luận và quyền được mua sắm vũ khí. Ngược lại, xã hội châu Á thường cố gắng cân bằng quyền của

cá nhân với các quyền của gia đình, tổ chức và chính phủ. Vì vậy, họ thường đưa ra những chính sách có tính cộng đồng hơn và ít tính cá nhân hơn. Những sự khác biệt như vậy thường khiến những người phương Tây phàn nàn về tính chuyên quyền của chính phủ các nước châu Á. Trong khi đó, người dân các nước châu Á thường phê bình xã hội phương Tây về tính ích kỷ, tỷ lệ tội phạm cao và tạo ra “sự tôn thờ chủ nghĩa cá nhân”. Mô thức của mỗi cá nhân như vậy giống như những “kính màu” định hướng quan điểm về thế giới và cách thức hình thành tư duy về các vấn đề của thế giới.

Thông thường, không dễ để có thể nhận diện các mô thức bởi vì chúng thường tiềm ẩn, không hiện hình và là giả định mà thôi. Tuy nhiên, việc nhận diện các mô thức này là chìa khóa để hiểu sự khác biệt trong nhận thức của con người về cùng một hiện tượng. Ví dụ, tuy cùng mục đích nhằm cải thiện chất lượng giáo dục phổ thông trung học, nhưng tại sao những người theo chủ nghĩa tự do lại tin rằng cách tốt nhất là sử dụng đội ngũ giáo viên giỏi; trong khi những người theo chủ nghĩa bảo thủ lại cho rằng nên tư nhân hóa ngành giáo dục, sử dụng nhà trường như một đơn vị kinh doanh? Bởi vì, những người bảo thủ đặt nhiều niềm tin hơn vào thị trường cạnh tranh (tự do cạnh tranh sẽ thúc đẩy chất lượng giáo dục), trong khi người theo khuynh hướng tự do lại tin vào sự lao động (nhiều giáo viên giỏi sẽ cải thiện chất lượng).

Tương tự trong nghiên cứu khoa học xã hội. Trong một nghiên cứu có nhiệm vụ giải thích tại sao một công nghệ nào đó được áp dụng rất thành công trong một công ty này nhưng lại thất bại thảm hại ở một nơi khác, thì nhà nghiên cứu nhìn thế giới thông qua một “thấu kính lý trí” sẽ tìm kiếm những giải thích có lý trí cho vấn đề đó, ví như công nghệ không hiệu quả hoặc không phù hợp với bối cảnh sử dụng. Trong khi đó, một nhà nghiên cứu khác sử dụng “thấu kính xã hội”, dù có nghiên cứu cùng một vấn đề nhưng lại đi tìm kiếm những khác biệt về mặt xã hội, ví như việc đào tạo, bố trí, sử dụng nhân sự không tốt hoặc thiếu sự hỗ trợ trong công tác quản lý. Còn nhà nghiên cứu sử dụng “thấu kính chính trị” thì sẽ tìm kiếm những vấn đề về chính trị trong tổ chức có thể ảnh hưởng đến tiến trình thực hiện công nghệ. Vì vậy, những mô thức trong tiềm thức sẽ ảnh hưởng đến các khái niệm, đến các quan sát và giải thích tiếp sau về một hiện tượng. Tuy vậy, nhiều hiện tượng xã hội có bản chất phức tạp, có thể các mô thức trên chỉ đúng một phần và để có được những hiểu biết toàn diện hơn đối với một hiện tượng xã hội đòi hỏi phải hiểu và áp dụng nhiều mô thức.

Có hai mô thức hiện thời được các nhà khoa học xã hội sử dụng phổ biến là chủ nghĩa thực chứng (positivism) và chủ nghĩa hậu thực chứng (post-positivism). **Chủ nghĩa**

thực chứng, dựa trên các công trình của nhà triết học người Pháp Auguste Comte (1798-1857), là mô thức thống trị trong khoa học cho tới giữa thế kỷ XX. Tinh thần của nó là khoa học hoặc sáng tạo tri thức nên bị giới hạn ở những gì mà con người có thể quan sát và đo lường được. Chủ nghĩa thực chứng có khuynh hướng chỉ tin tưởng vào các lý thuyết có thể kiểm chứng được trực tiếp. Mặc dù về nguồn gốc, nguyên lý này xuất phát từ chủ trương phân tách khoa học với tôn giáo (vì trong tôn giáo, người ta không thể kiểm chứng một cách khách quan những lời giáo huấn), nhưng thuyết thực chứng dẫn tới *chủ nghĩa thực nghiệm* (empiricism) hoặc dẫn tới sự tin tưởng tuyệt đối vào những dữ liệu đã quan sát được và chối bỏ những ý định mở rộng hay phân tích xa hơn các thực tế chưa hoặc không thể quan sát. Ví dụ do tư duy và xúc cảm của con người không thể cân, đo, đong, đếm trực tiếp, vậy nên, chúng không được xem là những đối tượng chính đáng của nghiên cứu khoa học.

Thất vọng với bản chất thực nghiệm chật hẹp của chủ nghĩa thực chứng đã dẫn đến sự hình thành **chủ nghĩa hậu thực chứng** (hay hậu hiện đại) trong thời gian nửa cuối thế kỷ XX. Chủ nghĩa hậu thực chứng cho rằng có thể đưa ra kết luận hợp lý về một hiện tượng bằng việc kết hợp các quan sát thực nghiệm với lập luận logic. Các nhà hậu thực chứng xem tri thức khoa học không phải là bất biến mà có tính chất xác xuất và biến thiên (nghĩa là tri thức có được dựa trên nghiên cứu nhiều hiện tượng ngẫu nhiên), việc khám phá các hiện tượng ngẫu nhiên sẽ giúp để hiểu thực tế xã hội tốt hơn. Sau này, những người theo thuyết hậu thực chứng phân chia thành *phái chủ nghĩa chủ quan* (subjectivist) và *phái hiện thực phê phán* (critical realist). Phái chủ nghĩa chủ quan cho rằng thế giới không phải là một thực tại khách quan mà là mô hình tưởng tượng trong ý thức chủ quan của con người. Phái hiện thực phê phán tin rằng có một thực thể khách quan tồn tại độc lập với ý thức chủ quan của con người, nhưng con người không bao giờ có thể nhận thức thấu đáo về thực thể đó.

Burrell và Morgan (1979), trong cuốn sách “*Các mô thức xã hội học và phân tích tổ chức*” (Sociological Paradigms and Organizational Analysis), đã đề xuất rằng phương pháp mà các nhà khoa học xã hội sử dụng để xem xét và nghiên cứu các hiện tượng xã hội bị chi phối bởi hai giả thiết cơ bản của triết học là bản thể luận (ontology) và nhận thức luận (epistemology). **Bản thể luận** phản ánh các giả thiết về cách nhìn nhận thế giới, xem thế giới là bản thể chứa đựng trật tự xã hội hay là thường xuyên thay đổi. **Nhận thức luận** phản ánh giả thiết về cách tốt nhất để nghiên cứu về thế giới, ví dụ như chúng ta nên sử dụng cách tiếp cận khách quan hay chủ quan để nghiên cứu hiện thực xã hội. Khi sử dụng

cả hai hệ thống giả thiết này, chúng ta có thể phân loại các nghiên cứu khoa học thành một trong bốn nhóm (xem Hình 3.1).

Nếu nhà nghiên cứu xem thế giới như một bản thể chứa đựng trật tự xã hội (bản thể luận) và do đó muốn tìm kiếm các mô hình nghiên cứu về các sự kiện và hành vi có trật tự. Họ tin rằng cách tốt nhất để nghiên cứu thế giới này là phải sử dụng cách tiếp cận khách quan (nhận thức luận). Hướng tiếp cận này không phụ thuộc vào người thực hiện các quan sát hay giải thích. Bằng việc sử dụng các công cụ thu thập dữ liệu tiêu chuẩn như các cuộc khảo sát thì đó là áp dụng mô thức **chức năng luận** (functionalism). Tuy nhiên, nếu họ tin rằng cách tốt nhất để nghiên cứu về trật tự xã hội là thông qua các suy luận chủ quan của những người tham gia, như bằng cách phỏng vấn những người tham gia khác nhau và sau đó sử dụng quan điểm chủ quan của chính nhà nghiên cứu để diễn giải sự khác nhau trong câu trả lời của người tham gia, thì như vậy họ đang theo đuổi mô thức **diễn giải** (interpretivism). Nếu người nghiên cứu tin rằng thế giới chứa đựng những thay đổi căn bản và muốn hiểu hoặc điều chỉnh những thay đổi đó bằng cách sử dụng cách tiếp cận khách quan, thì họ đang thực hành mô thức **cấu trúc luận cấp tiến** (radical structuralism). Nếu họ muốn hiểu những thay đổi trong xã hội bởi việc sử dụng quan điểm chủ quan của những người tham gia, thì họ đang đi theo mô thức **chủ nghĩa nhân văn cấp tiến** (radical humanism).



Hình 3.1. Bốn mô thức nghiên cứu khoa học xã hội
(Nguồn: Burrell and Morgan, 1979)

Cho đến nay, phần lớn các nghiên cứu khoa học xã hội áp dụng mô thức chức năng luận của khoa học tự nhiên. Những người theo mô thức này tin rằng trật tự xã hội hoặc các mô hình xã hội có thể được tìm hiểu thông qua các thành phần chức năng của chúng. Do đó, họ cố gắng phân chia một vấn đề thành các thành phần nhỏ, sau đó nghiên cứu chi tiết một hoặc một số thành phần nhỏ bằng cách sử dụng các phương pháp khách quan như khảo sát và nghiên cứu thực nghiệm. Tuy nhiên, với sự xuất hiện của thuyết hậu thực chứng, có một lượng nhỏ nhưng gia tăng nhanh các nhà nghiên cứu khoa học xã hội đang cố gắng tìm hiểu trật tự xã hội bằng cách sử dụng các phương pháp chủ quan chẳng hạn như phỏng vấn và nghiên cứu dân tộc học.

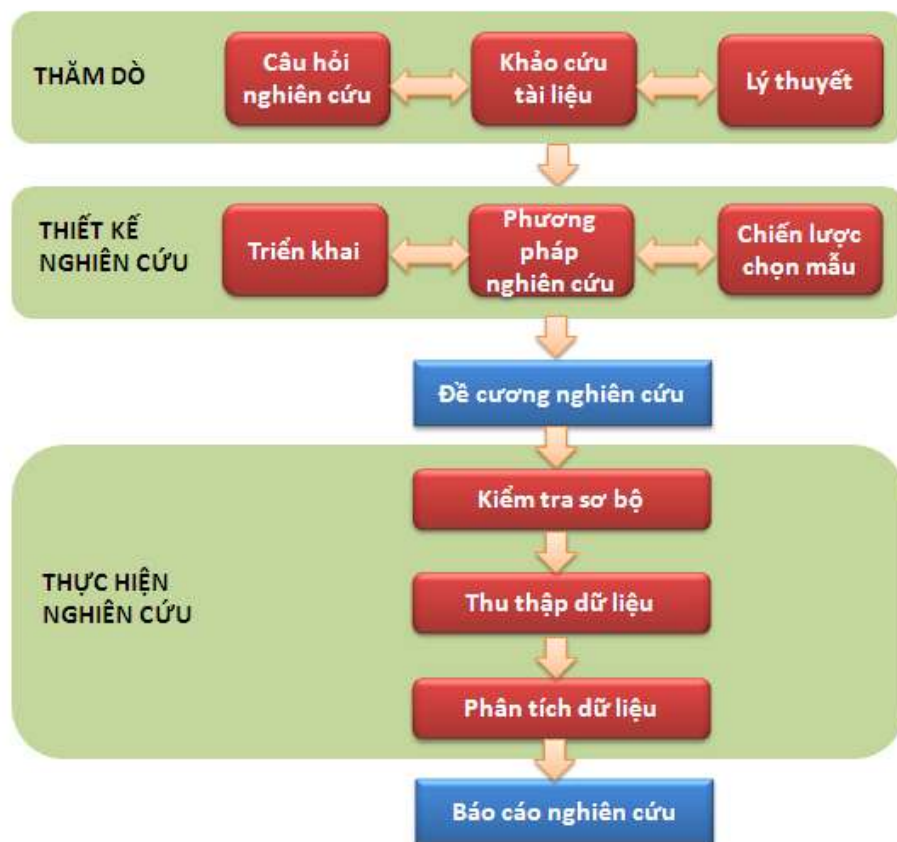
Chủ nghĩa nhân văn cấp tiến và cấu trúc luận cấp tiến tiếp tục chiếm một tỷ lệ nhỏ trong các nghiên cứu khoa học xã hội, bởi vì các nhà khoa học chủ yếu quan tâm đến sự hiểu biết về mô hình phổ quát của hành vi, sự kiện và hiện tượng, chứ không phải là những sự kiện mang đặc tính riêng hoặc đang thay đổi. Tuy nhiên, nếu bạn muốn nghiên cứu về thay đổi xã hội, chẳng hạn như lý do tại sao các phong trào dân chủ đang ngày càng nổi lên trong các quốc gia Trung Đông hoặc lý do tại sao phong trào này thành công ở Tunisia, nhưng mất khoảng thời gian dài hơn để thành công ở Libya và vẫn không thành công ở Syria. Vậy thì chủ nghĩa nhân văn cấp tiến có lẽ là cách tiếp cận đúng cho nghiên cứu này. Nhìn chung, các hiện tượng xã hội thường bao gồm các yếu tố của cả sự trật tự và sự thay đổi. Ví dụ, thành công của một tổ chức phụ thuộc vào quy trình kinh doanh, thủ tục vận hành và tính trách nhiệm trong công việc, đồng thời nó bị hạn chế bởi hàng loạt các yếu tố liên tục thay đổi trong thương trường như các đối thủ cạnh tranh, sản phẩm cạnh tranh, nhà cung cấp và khách hàng. Do đó, một sự hiểu biết toàn diện và đầy đủ hơn về hiện tượng xã hội như tại sao một số tổ chức thành công hơn những tổ chức khác, đòi hỏi sự cân nhắc để áp dụng tiếp cận đa mô thức trong nghiên cứu.

Tổng quan về tiến trình nghiên cứu

Vậy thì các mô thức tư duy của chúng ta tác động tới nghiên cứu khoa học xã hội như thế nào? Xét một cách bản chất nhất, tất cả các nghiên cứu khoa học là một quá trình lặp đi, lặp lại của sự quan sát (observation), biện giải (rationalization - lập luận giải thích) và kiểm nghiệm (validation). Trong giai đoạn **quan sát**, chúng ta quan sát một hiện tượng tự nhiên hoặc xã hội, sự kiện, hành vi mà ta quan tâm. Trong giai đoạn **biện giải**, chúng ta cố gắng hiểu các hiện tượng, sự kiện, hành vi đã quan sát bằng cách kết nối một cách logic các phần khác nhau của vấn đề cần giải quyết. Điều này, trong một số trường hợp, có thể dẫn đến xây dựng một lý thuyết. Cuối cùng, trong giai đoạn **kiểm nghiệm**, chúng ta kiểm

tra lý thuyết đã xây dựng bằng cách sử dụng các phương pháp khoa học thông qua quá trình thu thập và phân tích dữ liệu. Để làm được điều đó, có thể phải sửa đổi hoặc mở rộng lý thuyết ban đầu của chúng ta. Tuy nhiên, các thiết kế nghiên cứu có thể thay đổi tùy thuộc vào việc các nhà nghiên cứu bắt đầu bằng việc quan sát và cố gắng biện giải các quan sát (nghiên cứu quy nạp), hay nhà nghiên cứu bắt đầu với một biện giải hoặc một lý thuyết trước đó và cố gắng để kiểm nghiệm lý thuyết đó (nghiên cứu diễn dịch). Do đó, chu trình *quan sát - biện giải - kiểm nghiệm* là tương tự chu trình *quy nạp - diễn dịch* của nghiên cứu đã bàn đến trong Chương 1.

Nhìn chung, phần lớn các nghiên cứu truyền thống có xu hướng diễn dịch và chức năng luận. Hình 3.2 cung cấp sơ đồ của một dự án nghiên cứu như vậy. Hình này mô tả một loạt các hoạt động được thực hiện trong nghiên cứu chức năng luận, được phân loại thành ba giai đoạn: thăm dò, thiết kế và thực hiện nghiên cứu. Lưu ý rằng thiết kế khái quát này không phải là một sơ đồ hay lộ trình cho tất cả các nghiên cứu. Nó chỉ áp dụng cho các nghiên cứu chức năng luận. Bạn có thể và nên sửa đổi thiết kế này để phù hợp với yêu cầu của một nghiên cứu cụ thể.



Hình 3.2. Tiến trình nghiên cứu theo chức năng

Giai đoạn đầu của nghiên cứu là **thăm dò** (exploration). Giai đoạn này bao gồm việc tìm kiếm và lựa chọn các câu hỏi nghiên cứu cho những nghiên cứu sâu hơn, kiểm tra các công trình khoa học đã công bố trong cùng lĩnh vực nghiên cứu để hiểu được tình trạng tri thức khoa học hiện có và nhận diện các lý thuyết có thể giúp trả lời các câu hỏi nghiên cứu.

Nhiệm vụ của bước đầu tiên trong giai đoạn thăm dò là để xây dựng một hoặc nhiều câu hỏi nghiên cứu đối với một hành vi, sự kiện hay hiện tượng quan tâm. Câu hỏi nghiên cứu là những câu hỏi cụ thể mà bạn muốn tìm câu trả lời trong nghiên cứu của bạn. Ví dụ: Những yếu tố nào thúc đẩy người tiêu dùng mua hàng hóa và dịch vụ trực tuyến mà không cần biết ai là nhà cung cấp các hàng hóa, dịch vụ đó? Làm thế nào để có thể nâng cao tính sáng tạo trong học sinh trung học? Tại sao một số người lại thực hiện hành vi khủng bố? Câu hỏi nghiên cứu có thể đào sâu vào vấn đề như đó là cái gì, tại sao, làm thế nào, khi nào, v.v.. Câu hỏi nghiên cứu thú vị hơn là những những câu hỏi hấp dẫn với số đông quần chúng (ví dụ câu hỏi nghiên cứu: “các doanh nghiệp có thể đổi mới như thế nào” sẽ thú vị hơn câu hỏi "các công ty Trung Quốc có thể đổi mới như thế nào trong lĩnh vực dịch vụ"); nêu lên được những vấn đề thực tiễn phức tạp (ngược lại với các vấn đề vụn vặt mang tính giả thuyết) và chưa có các câu trả lời là rõ ràng. Các câu hỏi nghiên cứu có trọng tâm hẹp (ví dụ có câu trả lời nhị phân có/không) thường ít hữu ích, ít thú vị và chưa phù hợp để nắm bắt những sắc thái tinh tế của các hiện tượng xã hội. Các câu hỏi nghiên cứu không thú vị thường dẫn đến kết quả nghiên cứu không hấp dẫn và khó công bố.

Bước tiếp theo là tiến hành **khảo cứu tài liệu** (literature review) về lĩnh vực quan tâm. Khảo cứu tài liệu có ba mục đích là (1) để khảo sát trạng thái tri thức hiện có trong lĩnh vực nghiên cứu, (2) để xác định các tác giả, bài viết, lý thuyết và những kết quả nghiên cứu quan trọng trong lĩnh vực đó, và (3) để xác định khoảng trống tri thức trong lĩnh vực nghiên cứu. Ngày nay khảo cứu tài liệu thường được thực hiện bằng cách sử dụng các từ khóa trên máy tính để tìm kiếm trong cơ sở dữ liệu trực tuyến. Từ khóa có thể được kết hợp bằng cách sử dụng liên từ "and" và "or" để thu hẹp hoặc mở rộng các kết quả tìm kiếm. Ngay khi một danh sách ngắn các bài báo liên quan đến từ khóa được tìm thấy, người nghiên cứu phải đọc lướt qua mỗi bài viết hoặc ít nhất là phần tóm tắt của nó để xác định sự phù hợp của bài viết đó đối với lĩnh vực nghiên cứu, từ đó quyết định có đọc chi tiết hay không.

Khảo cứu tài liệu phải được đầu tư thời gian và công sức đúng mức, không nên chỉ giới hạn ở một vài bài báo khoa học, trong một vài năm hoặc một phương pháp luận cụ thể nào đó. Các bài viết đã nghiên cứu có thể được tóm tắt dưới dạng bảng biểu và có thể được

cấu trúc lại bằng cách sử dụng mô hình tổ chức như mô hình khái niệm. Quá trình khảo cứu tài liệu phải xác định xem các câu hỏi nghiên cứu ban đầu đã được đề cập trong các tài liệu đó hay chưa (điều này giúp loại trừ việc nghiên cứu trùng lặp); liệu có những câu hỏi nghiên cứu mới hơn, thú vị hơn; và liệu các câu hỏi nghiên cứu ban đầu có cần phải được sửa đổi sau những kết quả khảo cứu tài liệu. Quá trình khảo cứu tài liệu cũng có thể cung cấp một số câu trả lời trực giác hoặc câu hỏi tiềm năng cho các câu hỏi quan tâm và/ hoặc giúp xác định lý thuyết đã sử dụng trước đó để phát biểu các câu hỏi nghiên cứu tương tự.

Bởi vì nghiên cứu chức năng luận (diễn dịch) liên quan đến việc kiểm nghiệm lý thuyết, nên bước thứ ba là để xác định một hoặc nhiều lý thuyết có thể giúp phát biểu các câu hỏi nghiên cứu mong muốn. Trong quá trình khảo cứu tài liệu, có thể giúp phát hiện ra một loạt các khái niệm, phạm trù tiềm năng liên quan đến hiện tượng nghiên cứu; tìm được một lý thuyết phù hợp giúp xác định những phạm trù cụ thể nào thật sự liên quan và liên quan như thế nào đến hiện tượng nghiên cứu. Việc bỏ qua khâu lựa chọn lý thuyết phù hợp có thể dẫn đến hệ quả là chúng ta tìm kiếm, xây dựng những phạm trù ít hoặc không liên quan đến hiện tượng nghiên cứu; làm giảm cơ hội đạt được kết quả nghiên cứu có ý nghĩa bằng các phương pháp khoa học chứ không phải do may mắn hay ngẫu nhiên tìm thấy. Trong nghiên cứu chức năng luận, lý thuyết có thể được sử dụng làm cơ sở logic để xây dựng các giả thuyết định đề cho kiểm tra thực nghiệm. Rõ ràng, không phải toàn bộ các lý thuyết hiện có đều phù hợp cho nghiên cứu tất cả các hiện tượng xã hội. Vì vậy, các lý thuyết phải được lựa chọn cẩn thận, dựa trên tính tương thích giữa các tiền đề trong lý thuyết với các vấn đề nghiên cứu. Chúng ta sẽ tìm hiểu chi tiết các lý thuyết và quá trình xây dựng lý thuyết trong chương kế tiếp.

Giai đoạn thứ hai trong quá trình nghiên cứu là **thiết kế nghiên cứu** (research design). Quá trình này liên quan đến việc tạo ra một kế hoạch hoạt động để trả lời đầy đủ, thuyết phục các câu hỏi nghiên cứu đã xác định trong giai đoạn thăm dò. Điều này bao gồm việc lựa chọn một phương pháp nghiên cứu, thao tác hóa các phạm trù liên quan và xác định phương pháp lấy mẫu phù hợp.

Thao tác hóa (operationalization) là quá trình thiết kế các biện pháp đo lường, đánh giá chính xác đối với các phạm trù lý thuyết trừu tượng. Đây là một vấn đề quan trọng trong nghiên cứu khoa học xã hội. Việc đưa ra các phạm trù, chẳng hạn như sự thành kiến, tha hóa và chủ nghĩa tự do là rất khó để định nghĩa, chứ chưa nói tới việc đo lường chúng một cách chính xác, thấu đáo. Thao tác hóa bắt đầu với việc xác định rõ một “*định nghĩa thao tác*” hay còn gọi là “khái niệm hóa” (conceptualization) các phạm trù cần quan tâm.

Tiếp theo, nhà nghiên cứu có thể tìm kiếm các tài liệu để xem liệu có những khái niệm - công cụ đang còn giá trị, phù hợp với định nghĩa thao tác, có thể sử dụng trực tiếp hoặc có thể dùng để sửa đổi thành các phạm trù cần quan tâm. Nếu không có hoặc các khái niệm - công cụ hiện có không đủ hay phản ánh không đầy đủ phạm trù cần có, nhà nghiên cứu có thể xây dựng khái niệm - công cụ mới để đánh giá những phạm trù này. Công việc này giúp xác định chính xác làm thế nào có thể đánh giá các phạm trù mong muốn (ví dụ như có bao nhiêu mục, mục gì,...). Đây có thể là một quá trình lâu dài, đòi hỏi sự kiên trì, phải qua nhiều vòng kiểm định và sửa đổi trước khi các phương tiện mới thiết kế có thể được chấp nhận như là “giá trị khoa học”. Chúng ta sẽ bàn về hoạt động thao tác các phạm trù trong chương về đo lường tiếp theo.

Đồng thời với thao tác hoá, nhà nghiên cứu cũng phải quyết định lựa chọn **phương pháp nghiên cứu** để sử dụng cho việc thu thập dữ liệu giúp trả lời các câu hỏi nghiên cứu. Những phương pháp này có thể bao gồm các phương pháp định lượng như thử nghiệm hay thí nghiệm, nghiên cứu khảo sát; các phương pháp định tính như nghiên cứu trường hợp hay nghiên cứu hành vi hoặc có thể kết hợp cả định lượng và định tính. Nếu chọn thí nghiệm thì thiết kế thí nghiệm đó là gì? Nếu chọn khảo sát, bạn phải lập kế hoạch thực hiện cuộc khảo sát qua thư, điện thoại, internet hay kết hợp các cách thức đó? Đối với các hiện tượng xã hội phức tạp, đa dạng và không rõ ràng, các tiếp cận đa phương pháp có thể phù hợp hơn, giúp tận dụng những thế mạnh của mỗi phương pháp nghiên cứu và tạo ra kết quả mà việc sử dụng phương pháp đơn lẻ không thể có được.

Nhà nghiên cứu cũng phải cẩn thận lựa chọn khách thể nghiên cứu hay quần thể (population) mà họ muốn thu thập dữ liệu, một chiến lược **chọn mẫu** (sampling) để lấy mẫu từ những quần thể đó. Chẳng hạn, họ nên khảo sát các cá nhân hay các công ty hay các nhóm làm việc trong công ty? Những cá nhân hoặc các công ty nào mà họ muốn nhắm đến? Chiến lược chọn mẫu liên quan chặt chẽ với các đơn vị phân tích trong vấn đề nghiên cứu. Khi lựa chọn một mẫu, cần hết sức chú ý tránh việc lấy mẫu tùy tiện (ví dụ lấy mẫu dựa trên sự giản tiện), điều này dẫn đến sai số quan sát. Việc chọn mẫu sẽ được phân tích sâu trong chương sau.

Ở giai đoạn này đã hoàn toàn có cơ sở để viết **đề cương nghiên cứu** (research proposal), trong đó vạch ra chi tiết và lý do đằng sau tất cả các công việc cần thực hiện trong các giai đoạn nghiên cứu tiếp theo. Đề cương chi tiết phải nêu lên những câu hỏi nghiên cứu và lý do bạn muốn nghiên cứu, các nghiên cứu đã tiến hành trong cùng lĩnh vực, các lý thuyết sẽ được sử dụng cùng với các giả thuyết cần kiểm tra, cách tiến hành để

đánh giá các phạm trù; các phương pháp nghiên cứu dự kiến sử dụng và lý do sử dụng các phương pháp này, chiến lược chọn mẫu. Cơ quan, tổ chức tài trợ cho nghiên cứu thường đưa ra các yêu cầu như vậy để họ lựa chọn một đề cương tốt nhất. Ngay cả một dự án nghiên cứu không cần phải tìm kiếm kinh phí, đề cương nghiên cứu có thể xem như là một phương tiện hữu hiệu để tìm kiếm thông tin phản hồi từ các nhà nghiên cứu khác và xác định các rào cản tiềm tàng với dự án nghiên cứu (ví dụ liệu có một số phạm trù quan trọng chưa được đề cập trong nghiên cứu). Những phản hồi ban đầu này là rất có giá trị, bởi vì việc sửa chữa các sai sót trong thiết kế nghiên cứu là quá muộn nếu tiến hành sau khi đã thu thập dữ liệu nghiên cứu.

Sau khi đã xác định nghiên cứu ai (đối tượng nghiên cứu), công cụ nghiên cứu (khái niệm) và cách thức thu thập dữ liệu (phương pháp nghiên cứu), nhà nghiên cứu thực sự tiến tới giai đoạn thực hiện nghiên cứu. Giai đoạn này bao gồm kiểm tra thử nghiệm các dụng cụ đo lường, thu thập dữ liệu và phân tích dữ liệu.

Kiểm tra sơ bộ (pilot testing) là một công việc thường bị xem nhẹ nhưng hết sức cần thiết của quá trình nghiên cứu. Nó giúp phát hiện các lỗi tiềm tàng trong thiết kế nghiên cứu (ví dụ, liệu rằng những câu hỏi có dễ hiểu với người được hỏi hay không), cũng như độ tin cậy, tính khả thi và giá trị khoa học của các công cụ đánh giá. Các mẫu dùng cho thử nghiệm thường là một nhóm nhỏ trong số các đối tượng khảo sát. Sau khi thử nghiệm thành công, nhà nghiên cứu có thể tiến hành **thu thập dữ liệu** (data collection) đối với toàn bộ mẫu được lựa chọn từ trước. Các dữ liệu thu thập được có thể thiên về định lượng hoặc định tính, tùy thuộc phương pháp nghiên cứu được sử dụng.

Sau khi thu thập, dữ liệu được phân tích và giải thích nhằm mục đích rút ra kết luận cho những câu hỏi nghiên cứu. Tùy thuộc vào loại dữ liệu thu thập được (định tính hay định lượng), **phân tích dữ liệu** (data analysis) có thể sử dụng các phương pháp nghiên cứu định lượng (ví dụ, sử dụng các kỹ thuật thống kê như hồi quy hoặc mô hình phương trình cấu trúc) hoặc định tính (ví dụ mã hóa hoặc phân tích nội dung).

Giai đoạn cuối của nghiên cứu là việc chuẩn bị một **báo cáo nghiên cứu** (research report) để tổng kết lại toàn bộ quá trình nghiên cứu và các kết quả nghiên cứu với hình thức có thể là một bài báo nghiên cứu khoa học, một luận án hoặc chuyên khảo. Báo cáo này chỉ rõ một cách chi tiết tất cả các công việc đã thực hiện trong quá trình nghiên cứu (ví dụ, các lý thuyết đã sử dụng, các phạm trù được lựa chọn, các biện pháp đánh giá, phương pháp nghiên cứu, lấy mẫu đã sử dụng, v.v...) và lý do sử dụng, cũng như các kết quả từng

giai đoạn của quá trình nghiên cứu. Quá trình nghiên cứu phải được mô tả đầy đủ chi tiết để các nhà nghiên cứu khác để có thể lặp lại nghiên cứu của bạn, kiểm tra các kết quả nghiên cứu có được, đánh giá xem liệu những kết luận có được chấp nhận về mặt khoa học hay không. Tất nhiên, nếu đã có đề cương nghiên cứu chi tiết, sát thực thì quá trình viết báo cáo sẽ đơn giản và nhanh chóng hơn. Lưu ý rằng nghiên cứu sẽ không có giá trị trừ khi quá trình và kết quả nghiên cứu có thể được các thế hệ nghiên cứu tương lai kiểm định. Việc kiểm định này là yếu tố then chốt cho sự tiến bộ không ngừng của khoa học.

Các lỗi phổ biến trong nghiên cứu

Quá trình nghiên cứu luôn chứa đựng rất nhiều khó khăn và thử thách. Những người mới bước vào sự nghiệp nghiên cứu thường nhận thấy rằng, sau khi đầu tư lượng thời gian và công sức đáng kể vào một dự án nghiên cứu, câu hỏi nghiên cứu của họ không được trả lời một cách đầy đủ hoặc những phát hiện của họ không có giá trị hay nghiên cứu không có đủ luận cứ khoa học để “được chấp nhận”. Những vấn đề như vậy thường dẫn đến việc cơ quan khoa học từ chối các kết quả nghiên cứu của họ. Dưới đây là một số lỗi khá phổ biến.

Câu hỏi nghiên cứu thiếu hấp dẫn về mặt khoa học. Đôi khi, việc chọn những vấn đề "sở trường", thú vị đối với chúng ta nhưng lại không hấp dẫn với cộng đồng khoa học. Nghĩa là, chúng không tạo ra kiến thức mới hoặc cái nhìn sâu sắc về hiện tượng nghiên cứu. Bởi vì quá trình nghiên cứu đòi hỏi phải đầu tư rất lớn về thời gian và nỗ lực, nên nhà nghiên cứu phải khẳng định (và thuyết phục được người khác) rằng câu hỏi nghiên cứu họ đang tìm kiếm câu trả lời cho việc xử lý các vấn đề thực tiễn (không phải là giả thuyết), có ảnh hưởng đến một phần đáng kể dân số và chưa được giải quyết triệt để trong các nghiên cứu trước.

Theo đuổi một nghiên cứu nhất thời. Một sai lầm phổ biến khác là khi người nghiên cứu theo đuổi một chủ đề “hot”, mới nổi lên nhưng lại sẽ lạc hậu nhanh chóng. Ví dụ điển hình là các đề tài đang nghiên cứu các công nghệ phổ biến hiện nay. Bởi vì phải mất vài năm để các nghiên cứu hoàn thành và xuất bản, mối quan tâm nhất thời về những chủ đề đó có thể sẽ mất vào thời điểm hoàn thành nghiên cứu và gửi công bố. Một chiến lược tốt hơn có thể là chọn những chủ đề "vượt thời gian" luôn luôn tiếp tục tồn tại trong những năm tiếp theo.

Những vấn đề không thể nghiên cứu được. Một số vấn đề nghiên cứu dường như không thể được trả lời đầy đủ dựa trên bằng chứng đã quan sát cũng như dựa trên các phương pháp và trình tự được chấp nhận hiện nay. Tốt nhất nên tránh những vấn đề kiểu

như vậy. Tuy nhiên, cần chú ý rằng một số vấn đề ban đầu có thể là không thể nghiên cứu được bởi những định nghĩa mơ hồ, nhưng sau khi được sửa đổi hoặc điều chỉnh, chúng có thể trở thành những vấn đề hữu ích và có thể nghiên cứu được.

Ưu tiên sử dụng phương pháp nghiên cứu yêu thích. Nhiều nhà nghiên cứu có xu hướng cố tình viết lại vấn đề nghiên cứu để sử dụng các phương pháp nghiên cứu yêu thích của họ (ví dụ như nghiên cứu khảo sát). Đây là một xu hướng đầy rủi ro. Phương pháp nghiên cứu tốt nhất nên được lựa chọn phù hợp với vấn đề nghiên cứu, chứ không phải điều chỉnh vấn đề nghiên cứu cho phù hợp với phương pháp thường được nhà nghiên cứu sử dụng.

Khai thác dữ liệu ngược quy trình. Một số nhà nghiên cứu có kế hoạch tiến hành thu thập dữ liệu trước (bằng cách sử dụng các công cụ thu thập sẵn có), sau đó tìm cách xử lý cho phù hợp với dữ liệu đó. Cần phải lưu ý rằng, thu thập dữ liệu chỉ là một bước trong một quá trình lâu dài và phức tạp của việc lập kế hoạch, thiết kế và thực hiện nghiên cứu. Trong thực tế, một loạt các hoạt động khác cần thiết phải được tiến hành trước khi thu thập dữ liệu. Nếu nhà nghiên cứu tiến hành thu thập dữ liệu trước mà không có kế hoạch chi tiết thì các dữ liệu thu thập được có thể sẽ không phù hợp, có khiếm khuyết hoặc không có giá trị và những nỗ lực để thu thập dữ liệu của họ có thể trở nên hoàn toàn lãng phí. Sự dòi dào của dữ liệu đã thu thập được không thể bù đắp hết những thiếu hụt trong nhiệm vụ lên kế hoạch và thiết kế nghiên cứu, đặc biệt là việc thiếu các câu hỏi nghiên cứu thú vị.

Chương 4

LÝ THUYẾT TRONG NGHIÊN CỨU KHOA HỌC

Như đã trình bày từ chương trước, khoa học là tri thức được biểu hiện bằng một tập hợp các “lý thuyết”, được xây dựng bằng cách sử dụng các phương pháp khoa học. Trong chương này, chúng ta sẽ xem xét lý thuyết là gì, tại sao chúng ta cần các lý thuyết trong nghiên cứu, các thành phần của một lý thuyết là gì, làm thế nào để đánh giá lý thuyết, làm thế nào để áp dụng lý thuyết trong nghiên cứu và trình bày các ví dụ minh họa của năm lý thuyết thường xuyên được sử dụng trong nghiên cứu khoa học xã hội.

Lý thuyết

Lý thuyết là các giải thích về một hành vi, sự kiện hay hiện tượng tự nhiên hoặc xã hội. Chính thống hơn, lý thuyết khoa học là một hệ thống các phạm trù (khái niệm) và các luận điểm (mối quan hệ giữa những phạm trù) trình bày một giải thích hợp lý, có hệ thống và mạch lạc về một hiện tượng được quan tâm trong phạm vi một số giả thiết và điều kiện nhất định (Bacharach 1989)¹.

Lý thuyết nên giải thích lý do tại sao sự việc xảy ra hơn là chỉ đơn thuần mô tả hoặc dự đoán. Lưu ý rằng có thể dự đoán các sự kiện hoặc hành vi bằng cách sử dụng một tập hợp các yếu tố dự báo mà không nhất thiết phải giải thích lý do tại sao những sự kiện như vậy đang diễn ra. Ví dụ, các nhà phân tích thị trường dự đoán biến động của thị trường chứng khoán dựa trên thông báo thị trường, báo cáo thu nhập của các công ty lớn, các dữ liệu mới từ Cục Dự trữ Liên bang và các cơ quan khác, dựa trên mối tương quan đã quan sát trước đó. Dự đoán chỉ đơn thuần đòi hỏi phân tích mối tương quan (correlations). Ngược lại, giải thích yêu cầu sự hiểu biết về mối quan hệ nhân - quả (causations). Thiết lập mối quan hệ nhân quả đòi hỏi phải có ba điều kiện: (1) mối tương quan giữa hai phạm trù, (2) ưu tiên về thời gian (nguyên nhân phải diễn ra trước kết quả trong một khoảng thời gian nhất định), và (3) sự bác bỏ các giả thuyết thay thế (thông qua kiểm nghiệm). Các lý thuyết khoa học khác với các giải thích thần học, triết học,... ở chỗ lý thuyết khoa học có thể được kiểm nghiệm bằng cách sử dụng các phương pháp khoa học.

¹ Bacharach, S. B. (1989). “Organizational Theories: Some Criteria for Evaluation,” *Academy of Management Review* (14:4), 496-515.

Giải thích có thể là giải thích cá biệt hoặc giải thích phổ quát. **Giải thích cá biệt** (idiographic explanation) là giải thích duy nhất một tình huống một cách chi tiết và đặc thù. Ví dụ, bạn đã không làm tốt trong một kỳ thi vì: (1) bạn đã quên là bạn có một bài kiểm tra vào ngày hôm đó, (2) bạn đến muộn buổi thi do ùn tắc giao thông, (3) bạn hoảng sợ khi đang thi, (4) bạn phải làm việc muộn vào buổi tối hôm trước và không thể ôn thi hoặc thậm chí (5) con chó của bạn cắn nát cuốn sách mà bạn cần phải ôn tập. Những lời giải thích có thể chi tiết, chính xác, hợp lệ, nhưng chúng có thể không áp dụng cho các tình huống khác tương tự, thậm chí liên quan đến cùng một người, do đó không thể khái quát được. Ngược lại, **giải thích phổ quát** (nomothetic explanation) là việc tìm kiếm lời giải thích cho một nhóm các tình huống hoặc các sự kiện chứ không phải là một tình huống hoặc sự kiện cụ thể. Ví dụ, sinh viên làm kém trong các kỳ thi là do họ không dành đủ thời gian chuẩn bị cho các kỳ thi hoặc họ bị căng thẳng, thiếu tập trung hoặc bị một số rối loạn y khoa khác. Do trong giải thích phổ quát, các lý thuyết được thiết kế để có thể khái quát nhiều tình huống, nhiều sự kiện khác nhau nên có xu hướng ít chính xác, ít hoàn chỉnh và ít chi tiết hơn. Tuy nhiên, lý thuyết giải thích thường cô đọng, chỉ sử dụng ít các biến số giảng giải. Bởi mục đích của lý thuyết là đưa ra những giải thích có khả năng khái quát cho cơ chế hoạt động của các sự kiện, hành vi hoặc hiện tượng, nên nhìn chung, bản chất giải thích của lý thuyết là khái quát.

Khi tìm hiểu về lý thuyết, việc hiểu những gì không phải là lý thuyết cũng rất quan trọng. Lý thuyết không phải là dữ liệu, các sự kiện, mô hình hoạt động, phép phân loại hoặc các kết quả thực nghiệm. Một bộ sưu tập các sự kiện thực tế không phải là một lý thuyết, nó giống như một đồng gạch đá mà không phải là một ngôi nhà. Tương tự như vậy, một bộ sưu tập của các phạm trù (ví dụ, một hệ loại phạm trù) không phải là một lý thuyết. Lý thuyết phải đi vượt ra khỏi các phạm trù tới mức độ bao gồm cả các luận điểm, giải thích và điều kiện nhất định. Các dữ liệu, sự kiện hay kết quả tìm kiếm chỉ đơn thuần là những quan sát thực nghiệm, trong khi đó lý thuyết được xem xét ở mức độ nhận thức và dựa trên lập luận chứ không chỉ là những quan sát đơn thuần.

Có rất nhiều lợi ích khi sử dụng các lý thuyết vào trong nghiên cứu. *Đầu tiên*, lý thuyết cung cấp nhận thức cơ bản của sự xuất hiện các hiện tượng tự nhiên hoặc xã hội bằng cách giải thích đâu là nguyên nhân, điều kiện, quá trình vận động và kết quả vận động của hiện tượng đó. *Thứ hai*, lý thuyết cung cấp một hệ quy chiếu giúp đánh giá các kết quả nghiên cứu thực nghiệm trước; lý giải sự khác biệt kết quả nghiên cứu do các yếu tố ngẫu nhiên ảnh hưởng tới mối quan hệ giữa hai phạm trù trong các nghiên cứu khác. *Thứ ba*, các

lý thuyết cung cấp những định hướng cho các nghiên cứu trong tương lai bằng việc giúp nhận diện các phạm trù và các mối quan hệ có giá trị cho nghiên cứu sâu hơn. Thứ tư, lý thuyết có thể đóng góp vào kho tàng tri thức bằng cách lấp đầy khoảng trống giữa các lý thuyết khác nhau và tạo ra nhu cầu đánh giá lại các lý thuyết hiện có trong bối cảnh mới.

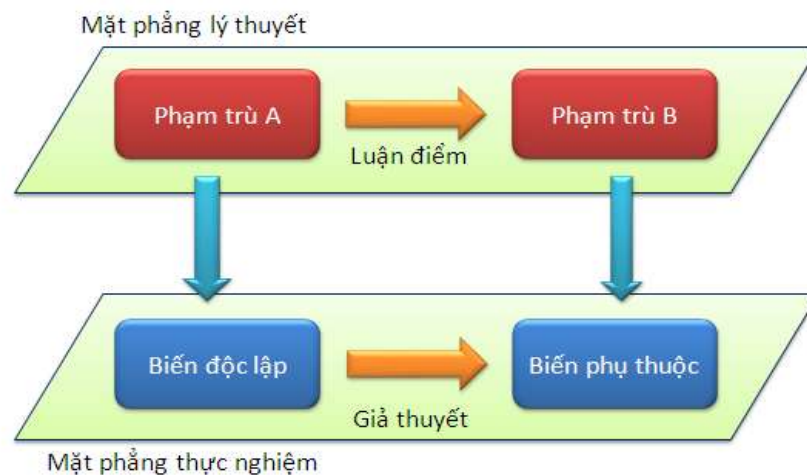
Tuy nhiên, lý thuyết cũng có thể có những hạn chế riêng nó. Lý thuyết không phải lúc nào cũng cung cấp các giải thích đầy đủ cho các hiện tượng quan tâm bởi nó chỉ giải thích đơn giản, khái quát thực tế bằng một hệ thống giới hạn các phạm trù và mối quan hệ. Lý thuyết được xây dựng để trở thành những giải thích đơn giản và ngắn gọn, trong khi thực tế có thể phức tạp hơn rất nhiều. Hơn nữa, lý thuyết có thể bị hạn chế bởi năng lực và tầm nhìn của nhà nghiên cứu, khiến họ bỏ lỡ các khái niệm quan trọng do chưa được định nghĩa trong lý thuyết.

Các thành tố của một lý thuyết

David Whetten (1989) đã đề xướng rằng có bốn yếu tố cấu thành nên một lý thuyết: phạm trù, luận điểm, lập luận và điều kiện hay giả thiết giới hạn. Phạm trù đảm nhiệm vai trò “là gì” của lý thuyết (ví dụ các khái niệm nào là quan trọng để giải thích một hiện tượng), luận điểm đảm nhiệm vai trò “như thế nào” (các khái niệm, phạm trù liên hệ với nhau như thế nào), lập luận giữ vai trò “tại sao” (ví dụ: tại sao các khái niệm liên kết với nhau như vậy) và điều kiện hay giả thiết giới hạn kiểm tra yếu tố “ai, khi nào và ở đâu” (như trong những tình huống cụ thể nào các phạm trù và mối quan hệ phát huy vai trò).

Phạm trù là khái niệm trừu tượng được khái quát ở mức độ cao, được lựa chọn cẩn thận để giải thích hiện tượng quan tâm. Nhớ lại ở Chương 2, phạm trù có thể là đơn (đại diện cho một khái niệm duy nhất) như độ tuổi, cân nặng; nhưng cũng có thể là phức (đại diện cho nhiều khái niệm ẩn) ví như tính cách cá nhân hay nền văn hóa. Trong khi một số phạm trù như độ tuổi, trình độ học vấn hay quy mô tổ chức có thể dễ hiểu, thì một số khác như tính sáng tạo, định kiến hay sự linh hoạt của tổ chức có thể phức tạp hơn; những phạm trù khác như sự tin tưởng, thái độ hay học tập có thể đề cập tới những xu thế nhất thời mà không phải là những trạng thái ổn định. Tuy vậy, tất cả các phạm trù phải có những định nghĩa thao tác rõ ràng, mạch lạc, chỉ ra một cách cụ thể cách thức các phạm trù đó được đo lường và cấp độ phân tích (cá nhân, nhóm hay tổ chức). Hình thức có thể đánh giá được của các phạm trù trừu tượng được gọi là các biến số. Ví dụ, chỉ số IQ là biến số được sử dụng để đánh giá một phạm trù trừu tượng đó là trí thông minh. Như đã đề cập ở phần trước, tiến trình nghiên cứu diễn ra trên hai mặt phẳng: một mặt phẳng lý thuyết và một

mặt phẳng thực nghiệm. Các phạm trù được nhận thức và định nghĩa ở mặt phẳng lý thuyết trong khi các biến được thao tác và đánh giá ở mặt phẳng thực nghiệm (bằng quan sát). Hơn nữa, các biến có thể độc lập, phụ thuộc, trung gian, điều hòa hay kiểm soát như được bàn luận trong Chương 2. Sự khác biệt giữa phạm trù và biến số được minh họa trong Hình 4.1.



Hình 4.1 Khác biệt giữa các khái niệm lý thuyết và thực nghiệm

Luận điểm là những mối liên hệ giữa các phạm trù dựa trên lập luận diễn dịch. Luận điểm được phát biểu dưới hình thức thông báo và chỉ rõ mối quan hệ nhân - quả (ví dụ, nếu X xuất hiện, sau đó Y sẽ hình thành). Lưu ý rằng luận điểm có thể là phỏng đoán, nhưng một điều bắt buộc là nội dung của nó có thể kiểm chứng được bằng quan sát thực nghiệm các biến tương ứng của chúng và nhà nghiên cứu nên loại bỏ những luận điểm không thỏa mãn điều kiện này. Tuy nhiên, giống như phạm trù, luận điểm được phát biểu ở mức độ lý thuyết, chúng chỉ có thể được kiểm định bằng cách kiểm tra các mối quan hệ giữa các biến tương ứng có thể đo lường được với những phạm trù đó. Việc hình thành các luận điểm qua thực nghiệm, thể hiện mối quan hệ giữa các biến, được gọi là giả thuyết. Sự khác biệt giữa các luận điểm (hình thành ở cấp độ lý thuyết) và giả thuyết (kiểm nghiệm ở cấp độ thực nghiệm) được mô tả trong Hình 4.1.

Cuối cùng, tất cả các lý thuyết đều bị giới hạn bởi các **giả định** (hay giả thiết - assumption) về giá trị, thời gian và không gian; **điều kiện biên** (boundary condition), các giả định và điều kiện này chỉ phối tình huống nào lý thuyết có thể được áp dụng và tình huống nào nó không có ý nghĩa. Ví dụ, nhiều lý thuyết kinh tế giả định rằng dù ít hay nhiều con người luôn thực dụng và duy lý. Chính vì vậy để hiểu hành vi con người, các lý thuyết

này theo đuổi cách tiếp cận tối đa hóa sự thiết thực, dựa trên kỳ vọng của người dùng về chi phí bỏ ra và lợi ích thu được. Ngược lại, các lý thuyết khoa học chính trị cho rằng con người đặt tham vọng chính trị lên trên lý trí, luôn cố gắng khẳng định mình cả trong công việc và cuộc sống cá nhân với nỗ lực tối đa hóa quyền lực của họ để kiểm soát những người khác. Xuất phát từ bản chất của các giả định cơ bản, các lý thuyết kinh tế và lý thuyết chính trị không thể so sánh trực tiếp, vì vậy nhà nghiên cứu không nên sử dụng các lý thuyết kinh tế nếu mục tiêu của họ là tìm hiểu cấu trúc quyền lực hay sự phân chia quyền lực trong một tổ chức. Tương tự như vậy, lý thuyết có thể có các giả định văn hóa ngầm (ví dụ, chúng áp dụng đối với văn hóa cá nhân hay tập thể), giả định thời gian (ví dụ, chúng áp dụng cho giai đoạn đầu hoặc giai đoạn cuối trong hành vi con người) và giả định không gian (ví dụ, chúng áp dụng đối với một số địa phương nhất định mà không áp dụng cho nơi khác). Để kiểm nghiệm và áp dụng đúng đắn một lý thuyết, tất cả các giả định tiềm ẩn của nó hình thành nên ranh giới của lý thuyết đó phải được nhìn nhận và hiểu rõ. Thật không may, trên thực tế các nhà lý thuyết hiếm khi phát biểu rõ ràng các giả định ẩn, dẫn đến việc áp dụng không đúng các lý thuyết với các vấn đề trong nghiên cứu.

Các yếu tố làm nên một lý thuyết tốt

Lý thuyết được đơn giản hóa và thường là những giải thích về một phần nào đó của thực tại xã hội phức tạp. Như vậy, những giải thích này có thể thuyết phục hoặc thiếu thuyết phục, do đó, có thể có những lý thuyết mạnh và lý thuyết yếu. Làm thế nào chúng ta có thể đánh giá sự mạnh - yếu của một lý thuyết? Các nhà nghiên cứu đã đề xuất nhiều tiêu chí khác nhau, trong số đó có các tiêu chí quan trọng được liệt kê dưới đây:

- **Lập luận nhất quán** (logical consistency): Liệu các phạm trù, luận điểm, điều kiện biên và giả định về lý thuyết có được liên kết một cách thống nhất với nhau hay không? Trong một lý thuyết, nếu một vài trong số những thành tố này không nhất quán với nhau, ví dụ, một giả định lý thuyết tiếp cận theo hướng duy lý, nhưng một số phạm trù lại đại diện cho các khái niệm không duy lý. Vậy thì nó không phải là một lý thuyết mạnh.

- **Năng lực giải thích** (explanatory power): Lý thuyết đó giải thích (hoặc dự đoán) về hiện tượng nghiên cứu ở mức độ như thế nào? Các lý thuyết tốt giải thích rõ ràng, toàn diện, chính xác khách thể nghiên cứu, ví dụ chúng có thể được đánh giá bởi giá trị phương sai (R-square) trong các phương trình hồi quy.

- **Tính phản nghiệm** (falsifiability): trong thập niên 1940, nhà triết học người Anh Karl Popper đã khẳng định rằng lý thuyết muốn có giá trị, chúng phải được kiểm chứng.

Tính phản nghiệm mang đến tình huống lý thuyết hoàn toàn có thể bị bác bỏ nếu dữ liệu thực nghiệm không phù hợp với các luận điểm lý thuyết. Nói cách khác, lý thuyết không còn là lý thuyết trừ khi chúng có thể được kiểm định bằng thực nghiệm. Những phát biểu lặp lại như "một ngày có nhiệt độ cao là một ngày nóng" không cần phải kiểm nghiệm bằng thực nghiệm, bởi vì hiển nhiên một ngày nóng được xác định (và đo được mức nhiệt) là một ngày với nhiệt độ cao. Do đó, phát biểu như vậy không thể được xem như là một luận điểm lý thuyết. Tính phản nghiệm đòi hỏi các lý thuyết phải đưa ra các giải thích thay thế khác nhau để đảm bảo các phạm trù của nó được đánh giá một cách thấu đáo. Tuy nhiên, cần lưu ý khi nói rằng một lý thuyết có tính phản nghiệm không có nghĩa là lý thuyết đó nên bị làm sai lệch. Nếu một lý thuyết thực sự bị làm sai lệch dựa trên bằng chứng thực nghiệm, là một lý thuyết yếu!

- **Tính tối giản** (parsimony): sự cô đọng đề cập đến số lượng biến được sử dụng để giải thích một hiện tượng. Khái niệm này được nhà logic học người Anh William of Ockham đưa ra từ thế kỷ XIV (và được gọi là nguyên lý “dao cạo Ockham” - "Ockham's razor"). Ockham tuyên bố rằng trong số các lý thuyết giải thích về một hiện tượng quan sát thì lý thuyết nào đơn giản nhất (tức là sử dụng ít các biến nhất hoặc đưa ra ít giả định nhất) là lý thuyết tốt nhất. Để giải thích về một hiện tượng xã hội phức tạp, lý thuyết thường sử dụng ngày càng nhiều các phạm trù. Điều này mâu thuẫn với phương châm đơn giản hóa và khái quát hóa trong quá trình hình thành lý thuyết. Sự tối giản ảnh hưởng đến mức độ linh động của một lý thuyết. Các lý thuyết tối giản mang đến sự linh động ở mức độ cao hơn, cho phép chúng được khái quát hóa dễ dàng hơn cho các bối cảnh khác, môi trường khác và đối tượng khác.

Cách tiếp cận xây dựng lý thuyết

Làm thế nào để các nhà nghiên cứu xây dựng lên các lý thuyết? Steinfeld và Fulk (1990)² đề xuất bốn cách tiếp cận cho vấn đề này. *Cách tiếp cận đầu tiên* là xây dựng lý thuyết theo lối quy nạp dựa trên các mô hình về sự kiện hoặc hành vi đã quan sát. Cách tiếp cận này thường được gọi là “thiết lập lý thuyết”, bởi vì lý thuyết được thiết lập bởi các quan sát thực nghiệm. Kỹ thuật này phụ thuộc rất nhiều vào khả năng quan sát và giải thích của nhà nghiên cứu. Do vậy, lý thuyết được thiết lập có thể mang tính chủ quan và khó có thể được xác thực. Hơn nữa, việc quan sát một số mô hình sự kiện nhất định không phải lúc nào cũng có thể tạo ra lý thuyết, trừ khi nhà nghiên cứu có thể đưa ra lời giải thích phù

² Steinfeld, C.W. and Fulk, J. (1990). "The Theory Imperative," in *Organizations and Communications Technology*, J. Fulk and C. W. Steinfeld (eds.), Newbury Park, CA: Sage Publications.

hợp cho các mô hình quan sát. Chúng ta sẽ thảo luận về hướng tiếp cận thiết lập lý thuyết trong chương tiếp sau về nghiên cứu định tính.

Cách tiếp cận thứ hai để xây dựng lý thuyết là dựa trên một mô hình nhận thức đã xác định từ trước, thực hiện phân tích toàn diện hiện tượng nghiên cứu để nhận diện đặc điểm của hiện tượng đó. Ví dụ về mô hình nhận thức này là mô hình đầu ra - đầu vào đơn giản. Mô hình này giúp nhà nghiên cứu có thể tìm kiếm các hệ loại đầu vào khác nhau, ví dụ như những yếu tố cá nhân, tổ chức và công nghệ có tiềm năng liên quan tới hiện tượng nghiên cứu (đầu ra) và mô tả quá trình liên kết các yếu tố đó với hiện tượng nghiên cứu. Đây cũng là một cách tiếp cận quy nạp, nó phụ thuộc nhiều vào khả năng tổng hợp của người nghiên cứu; và do đó việc xây dựng lý thuyết có thể bị tác động bởi những định kiến trước đó của người nghiên cứu về hiện tượng nghiên cứu.

Cách tiếp cận thứ ba là mở rộng hoặc điều chỉnh các lý thuyết hiện có nhằm mục đích giải thích trong một bối cảnh mới, ví dụ mở rộng các lý thuyết về học tập của cá nhân để giải thích việc học tập của tổ chức. Khi mở rộng lý thuyết, một vài khái niệm, luận điểm và điều kiện biên của lý thuyết trước đó có thể vẫn được giữ lại; một số khác cần được điều chỉnh cho phù hợp với bối cảnh mới. Hướng tiếp cận diễn dịch này là một phương pháp hiệu quả để phát triển các lý thuyết hiện có, góp phần làm giàu tri thức nhân loại.

Cách tiếp cận thứ tư là áp dụng các lý thuyết hiện có trong những bối cảnh hoàn toàn mới bằng cách chỉ rõ những điểm khác biệt cơ bản giữa các bối cảnh khác nhau. Lối tiếp cận này dựa trên sự so sánh giống và khác nhau. Đây có thể là phương pháp hình thành lý thuyết sáng tạo nhất trong hướng tiếp cận diễn giải. Ví dụ Markus (1987)³ đã sử dụng phép so sánh loại suy giữa một vụ nổ hạt nhân và sự phát triển tự do của các mạng lưới (còn gọi là mạng lưới kinh doanh đa cấp) để đưa ra lý thuyết về lượng tối thiểu của sự phát triển mạng lưới đa cấp. Giống một vụ nổ hạt nhân đòi hỏi một lượng tới hạn các chất phóng xạ để duy trì sự nổ, Markus nhận thấy một mạng lưới đa cấp cũng đòi hỏi một số lượng tối thiểu người dùng để duy trì sự phát triển của nó, nếu không có lượng tối thiểu này, người dùng có thể bỏ mạng lưới đó, cuối cùng dẫn đến sự sụp đổ của mạng.

Ví dụ về các lý thuyết khoa học xã hội

Trong phần này, chúng tôi trình bày tổng quan ngắn gọn về một vài lý thuyết minh họa từ các bộ môn khoa học xã hội khác nhau. Những lý thuyết giải thích các hành vi xã

³ Markus, M. L. (1987). "Toward a 'Critical Mass' Theory of Interactive Media: Universal Access, Interdependence, and Diffusion," *Communication Research* (14:5), 491-511

hội loại khác nhau qua việc sử dụng một tập hợp các phạm trù, mệnh đề, điều kiện biên, giả định và logic nền tảng. Lưu ý rằng dưới đây chỉ là giới thiệu sơ lược về các lý thuyết, độc giả nên xem thêm các tài liệu gốc về những lý thuyết này để biết thêm nội dung chi tiết trong mỗi lý thuyết.

Lý thuyết về đại diện (Agency theory)^(*). Lý thuyết về đại diện (còn gọi là lý thuyết giữa người chủ và người đại diện), một lý thuyết cổ điển trong kinh tế học tổ chức, do Ross (1973)⁴ là người đầu tiên đề xướng. Lý thuyết này nhằm giải thích mối quan hệ kinh tế hai bên (như giữa chủ và người làm công, giữa giám đốc điều hành và cổ đông, giữa người mua và người bán) có mục tiêu không đồng nhất với nhau. Mục đích của lý thuyết về đại diện là để xác định rõ các hợp đồng và các điều kiện tối ưu thực hiện hợp đồng nhằm giảm thiểu hậu quả xấu xảy ra. Dựa trên giả thiết cốt lõi rằng con người có bản tính tư lợi và sợ rủi ro, lý thuyết này có thể được áp dụng ở cấp độ cá nhân hoặc tổ chức

Hai bên trong lý thuyết này là người chủ và người đại diện quản lý; người chủ thuê người đại diện thay mặt của mình để thực hiện một số nhiệm vụ. Trong khi mục tiêu của người chủ đòi hỏi hoàn thành nhanh, có hiệu quả các nhiệm vụ được giao, thì mục tiêu của người đại diện là làm việc theo tiến độ riêng của mình, tránh các rủi ro và tìm kiếm lợi ích cá nhân (như thu nhập cá nhân) trong số lợi ích của công ty. Do đó, các mục tiêu của họ không tương thích.

Bản chất của vấn đề có thể là do thiếu thông tin giữa hai bên. Do người chủ không có đầy đủ thông tin về phẩm chất đạo đức của người đại diện hoặc không đủ thông tin để đánh giá chính xác các kỹ năng của người đại diện. Việc thiếu thông tin như vậy có thể dẫn đến hệ quả là người đại diện không nỗ lực để hoàn thành tốt công việc được giao (rủi ro đạo đức); hoặc không hoàn thành công việc do thiếu kiến thức chuyên môn hoặc thiếu kỹ năng giải quyết công việc (rủi ro do lựa chọn). Các hợp đồng thông thường dựa trên hành vi (thời gian làm việc) để trả lương hàng tháng và như vậy không thể giải quyết được vấn đề này. Do vậy, *lý thuyết về đại diện* khuyến cáo xây dựng hợp đồng lao động dựa trên kết quả làm việc (đầu ra), chẳng hạn như tiền hoa hồng hoặc tiền thưởng được trả dựa trên

(*) Theo Lý thuyết về đại diện (thuyết đại diện), quan hệ giữa các cổ đông và người quản lý công ty được hiểu như là quan hệ đại diện - hay quan hệ ủy thác theo hợp đồng. Theo đó các cổ đông (những người chủ - principals), bổ nhiệm, chỉ định người khác (agents) để thực hiện quản lý công ty cho họ. Theo luật về công ty phổ biến trên thế giới, các cổ đông (đại hội đồng cổ đông) có quyền chọn lựa để bầu, bổ nhiệm các vị trí quản lý quan trọng của công ty như hội đồng quản trị hay tổng giám đốc (giám đốc điều hành), trao thẩm quyền ra các quyết định nhất định (theo điều lệ công ty) để hành động cho công ty và vì công ty, cũng như định đoạt tài sản của công ty (ND).

⁴ Ross, S. A. (1973). "The Economic Theory of Agency: The Principal's Problem," American Economic Review (63:2), 134-139.

mức độ hoàn thành công việc. Hoặc có thể ký kết hợp đồng hỗn hợp vừa dựa trên hành vi (thời gian làm việc), vừa dựa vào kết quả công việc. Các hợp đồng ủy quyền mua bán chứng khoán cho nhân viên là ví dụ về loại hợp đồng dựa trên kết quả, trong khi lương của nhân viên là một hợp đồng dựa trên hành vi.

Lý thuyết về đại diện cũng khuyến cáo một số công cụ mà người chủ có thể sử dụng để nâng cao hiệu quả của hợp đồng dựa trên hành vi, chẳng hạn như đầu tư cải thiện cơ chế giám sát (như thuê giám sát viên) để tránh tình trạng thông tin không đầy đủ; xây dựng các hợp đồng có thời hạn, việc gia hạn phụ thuộc vào hiệu quả hoạt động của các đại lý hoặc bằng cách điều chỉnh tiêu chí mô tả công việc sao cho thuận lợi hơn cho quá trình đánh giá hiệu quả công việc đã giao.

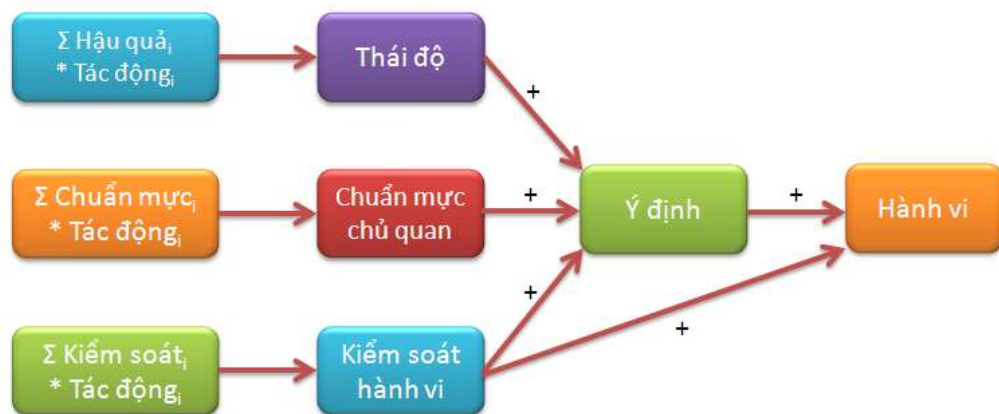
Lý thuyết hành vi hoạch định (Theory of Planned Behavior - Lý thuyết hành động theo dự tính/theo kế hoạch). Lý thuyết hành vi hoạch định do Ajzen (1991)⁵ đề xướng, là một lý thuyết tổng quát của hành vi con người trong bộ môn tâm lý xã hội, có thể được sử dụng để nghiên cứu một loạt các hành vi cá nhân. Nó giả định rằng hành vi của cá nhân là kết quả của quá trình lựa chọn có ý thức, bị chi phối bởi năng lực nhận thức cá nhân và áp lực xã hội. Lý thuyết này cho rằng hành vi cá nhân bị chi phối bởi kế hoạch hành động của họ trong một tình huống cụ thể, kế hoạch này sẽ tác động đến thái độ của người đó đối với tình huống, chuẩn mực chủ quan (subjective norm) và cách thức kiểm soát tình huống đó (xem hình 4.2).

Thái độ được định nghĩa là tổng thể những cảm xúc tích cực hay tiêu cực của cá nhân khi quan sát diễn biến của tình huống. Do đó, thái độ có thể được xem như là một tổng thể cảm nhận của một người về những hậu quả khác nhau của tình huống đó; và như vậy thái độ có thể được đo lường bằng cường độ của những hậu quả này. Quan niệm chủ quan liên quan đến nhận thức cá nhân về sự mong đợi của người thân về cách ứng xử của họ trong tình huống đó. Quan niệm chủ quan có thể được đo lường bằng một sự kết hợp cộng hưởng các nhận định của họ về mong muốn của các nhóm có liên quan như bạn bè, đồng nghiệp, người quản lý của họ. Kiểm soát hành vi là nhận thức cá nhân về các kiểm soát bên trong hoặc bên ngoài chi phối hành vi trong tình huống cụ thể. Kiểm soát bên trong có thể bao gồm năng lực cá nhân để thực hiện hành vi dự định (tự tác động), còn kiểm soát bên ngoài đề cập đến các nguồn lực bên ngoài sẵn có, cần thiết để thực hiện hành vi (điều kiện thuận lợi). Lý thuyết hành vi hoạch định cũng chỉ ra rằng đôi khi người ta có thể có ý định thực

⁵ Ajzen, I. (1991). "The Theory of Planned Behavior," *Organizational Behavior and Human Decision Processes* (50), 179-211.

hiện một hành vi nhất định nhưng thiếu các điều kiện cần thiết để thực hiện; và do đó lý thuyết này thừa nhận rằng yếu tố kiểm soát hành vi có thể mang đến những tác động trực tiếp tới diễn biến của hành vi, kết hợp với các tác động gián tiếp tạo ra bởi kế hoạch hành động của họ.

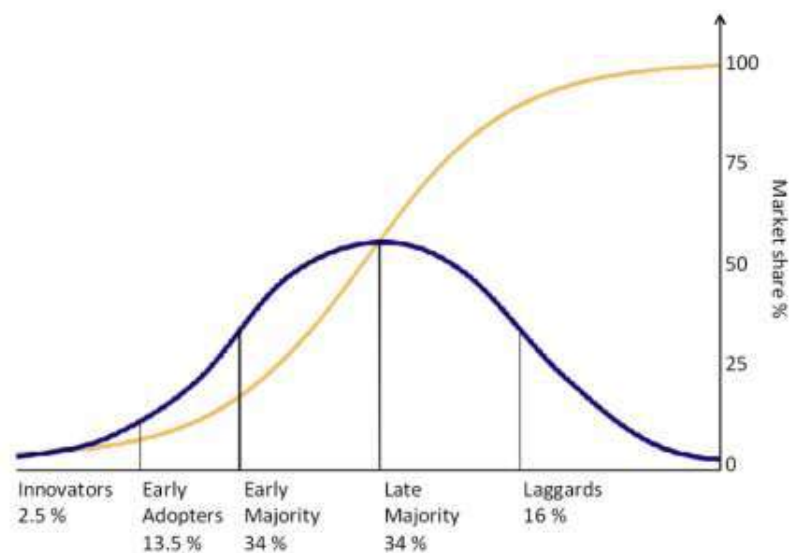
Lý thuyết hành vi hoạch định là một phần mở rộng của một lý thuyết trước đó được gọi là Lý thuyết hành động hợp lý, nó bao gồm thái độ và quan niệm chủ quan là yếu tố chính chính phối ý định (kế hoạch hành động), nhưng không chứa đựng yếu tố kiểm soát hành vi. Yếu tố kiểm soát hành vi được Ajzen bổ sung để giải thích cho trường hợp khi cá nhân thiếu một số điều kiện cần thiết để thực hiện kế hoạch hành động của mình (chẳng hạn như không có truy cập Internet tốc độ cao để lướt web).



Hình 4.2. Lý thuyết hành vi hoạch định

Lý thuyết khuyến tán đổi mới (Innovation Diffusion Theory - IDT). Lý thuyết khuyến tán đổi mới là một lý thuyết quan trọng trong chuyên ngành truyền thông học, nhằm mục đích giải thích các cơ chế lan truyền và tối đa hóa những ý tưởng cải tiến, nhân rộng các mô hình đổi mới diễn ra trong một cộng đồng. Các khái niệm cơ bản trong lý thuyết được nhà xã hội học Pháp Gabriel Tarde đưa ra lần đầu tiên. Tuy nhiên, lý thuyết này được phát triển hoàn chỉnh năm 1962 bởi Everett Rogers dựa trên kết quả của 508 nghiên cứu về khuyến tán. Bốn thành tố trụ cột của lý thuyết gồm: sự đổi mới hay các cải tiến, kênh truyền thông, thời gian và cấu trúc xã hội. Sự đổi mới có thể bao gồm công nghệ mới, hành vi mới hoặc ý tưởng mới. Chủ thể áp dụng những đổi mới này có thể ở cấp độ cá nhân hoặc tổ chức.

Ở cấp độ vĩ mô (trong một cộng đồng, một tổ chức), IDT nhìn nhận khuyến tán đổi mới là một tiến trình giao tiếp, trong đó các cá nhân trong một cộng đồng với đặc điểm cấu trúc nhất định tìm hiểu về một mô hình đổi mới và lợi ích của nó thông qua các kênh giao tiếp khác nhau (ví dụ như thông qua các phương tiện truyền thông hoặc thông qua những cá nhân đã và đang áp dụng mô hình này), sau khi bị thuyết phục bởi những lợi ích của mô hình, đã quyết định áp dụng nó. Khuyến tán là một tiến trình đòi hỏi thời gian, bắt đầu khá chậm rãi với một số ít người sử dụng, sau đó được lan truyền nhanh hơn khi nó trở thành xu thế trong cộng đồng và cuối cùng nó diễn ra chậm lại khi đạt đến sự bão hòa về số người sử dụng trong cộng đồng đó. Như vậy, ở cấp độ vĩ mô này, đồ thị tổng quát mô tả quá trình sử dụng các mô hình đổi mới là một đường cong chữ S như được hiển thị trong Hình 4.3 và đồ thị phân phối người sử dụng là một phân phối thường. Điều này có nghĩa là các chủ thể áp dụng là không giống nhau.



Hình 4.3. Đường cong khuyến tán hình chữ S

Tuy vậy, căn cứ vào thời điểm áp dụng các mô hình cải tiến hay đổi mới, họ có thể được phân thành người đổi mới (số lượng rất ít), một nhóm nhỏ người áp dụng ngay ban đầu, nhóm tương đối lớn áp dụng sớm, nhóm tương đối lớn áp dụng muộn và nhóm nhỏ người áp dụng muộn. Tỷ lệ số người áp dụng một mô hình đổi mới cũng phụ thuộc vào đặc điểm cấu trúc xã hội của cộng đồng đó. Ví dụ như có hay không sự hiện diện của các lãnh đạo tinh thần (những người mà quan điểm của họ được đánh giá cao) và các cá nhân có ảnh hưởng lớn đến hành vi của những người khác.

Ở cấp độ vi mô (cấp độ cá nhân), Rogers (1995)⁶ chỉ ra rằng quá trình áp dụng một mô hình đổi mới ở một cá nhân trải qua năm giai đoạn: (1) tìm hiểu: mỗi cá nhân tìm hiểu thông tin về mô hình mới thông qua các phương tiện truyền thông đại chúng hoặc các kênh giao tiếp cá nhân, (2) thuyết phục: họ bị thuyết phục bởi những tiện lợi hoặc bất lợi của mô hình, (3) quyết định: quyết định áp dụng hoặc từ chối áp dụng mô hình, (4) thực hiện: bước đầu thử nghiệm mô hình, và (5) xác nhận: quyết định có hay không áp dụng mô hình ở mức độ tối đa (Xem Hình 4.4). Có năm yếu tố ảnh hưởng đến các quyết định này của chủ thể áp dụng: (1) mức độ tiện lợi: những tiện lợi của mô hình mới so với các mô hình trước đó, (2) khả năng tương thích: mức độ phù hợp với thói quen, sở trường, quan niệm của người áp dụng, (3) độ phức tạp: mô hình có dễ dàng để hiểu và thực hiện hay không, (4) tính thử nghiệm: mô hình có thể được áp dụng thử nghiệm mức độ nào, và (3) khả năng kiểm nghiệm: các kết quả của việc áp dụng mô hình đổi mới được cảm nhận rõ ràng đến mức độ nào. Hai yếu tố cuối cùng mà Rogers đề xuất đã bị lược bỏ trong các nghiên cứu sau đó về đổi mới. Trong số năm yếu tố này, độ phức tạp sẽ kìm hãm việc áp dụng các mô hình đổi mới, trong khi đó bốn yếu tố còn lại thúc đẩy việc áp dụng này.

Ngoài ra, việc áp dụng đổi mới cũng phụ thuộc vào đặc điểm cá nhân, ví dụ như tính ưa mạo hiểm, trình độ học vấn, sự cởi mở và thói quen giao tiếp. Người áp dụng sự đổi mới ngay lúc ban đầu thường là những người ưa mạo hiểm, có học vấn và tìm hiểu phần lớn thông tin về mô hình đổi mới trên các phương tiện truyền thông đại chúng. Trong khi đó, những người áp dụng mô hình muộn hơn thường dựa trên các quan hệ cá nhân như người thân hoặc bạn bè để tìm hiểu những thông tin quan trọng về mô hình đổi mới. Bên cạnh những thành công, IDT đã gặp phải những chỉ trích về sự thiên vị dành cho xu hướng đổi mới, bởi vì nó giả định rằng tất cả sự đổi mới là có lợi và đến một thời điểm nào đó chúng sẽ trở nên thịnh hành trong cộng đồng. Ngoài ra, lý thuyết này bị chỉ trích vì nó cũng làm chậm quá trình loại bỏ những cải tiến hay các mô hình đổi mới không phù hợp với cộng đồng để thay thế bằng những mô hình phù hợp hơn, thiết thực hơn.

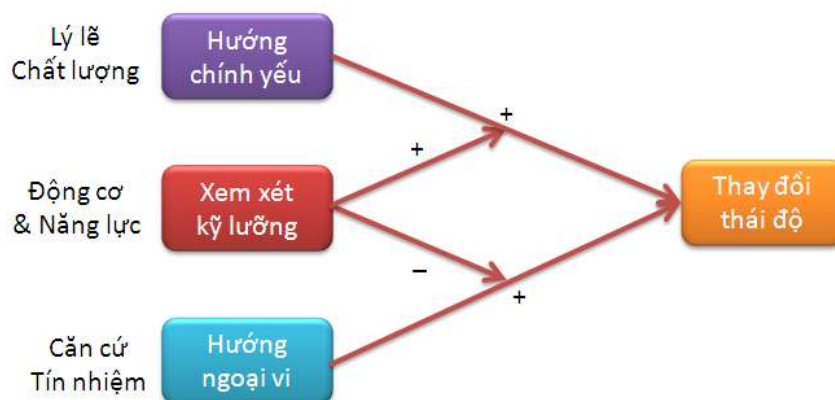


Hình 4.4. Quá trình áp dụng đổi mới

⁶ Rogers, E. (1962). Diffusion of Innovations. New York: The Free Press. Other editions 1983, 1996, 2005.

Mô hình xem xét kỹ lưỡng (Elaboration Likelihood Model - ELM). Mô hình xem xét kỹ lưỡng được Petty và Cacioppo (1986)⁷ phát triển, là lý thuyết trong tâm lý học về một tiến trình kép trong việc hình thành và thay đổi thái độ cá nhân. Lý thuyết này giải thích quá trình các cá nhân thay đổi quan điểm, thái độ của họ về một sự vật, sự kiện hoặc một hành vi sau quá trình xem xét của họ về sự vật, hiện tượng đó.

Mô hình xem xét kỹ lưỡng cho rằng quan điểm, thái độ của một cá nhân được phân thành hai “hướng” ảnh hưởng: hướng chính yếu (central route) và hướng ngoại vi (peripheral route). Hai hướng này khác nhau ở mức độ chuyên sâu trong quá trình phân tích, xử lý thông tin (xem Hình 4.5). Trong hướng chính yếu, chủ thể sẽ xem xét một cách chi tiết, đầy đủ các thông tin khác nhau về một hiện tượng, đánh giá mức độ liên quan và tính khách quan, toàn diện của các nguồn thông tin đó trước khi thận trọng đưa ra phán xét của mình về hiện tượng quan tâm. Nói cách khác, chủ thể với hướng chính yếu chú ý đến chất lượng nguồn thông tin. Trong khi đó, đối với hướng ngoại vi, ví dụ để hình thành quan điểm, thái độ về một mặt hàng, người tiêu dùng dựa vào các “tín hiệu” bên ngoài, ví dụ như số lượng người đã và đang dùng sản phẩm này, số lượng khuyến nghị của các chuyên gia hoặc số lượng người ủng hộ sản phẩm đó, chứ không phải dựa trên đặc tính của các nguồn thông tin này. Có thể nói hướng ngoại vi đòi hỏi sự đầu tư về trí tuệ ít hơn. Trong các nghiên cứu sử dụng mô hình xem xét kỹ lưỡng, các hướng nhằm thay đổi thái độ thường được xem xét thông qua hai phạm trù *đặc tính chính yếu* và *tín hiệu ngoại vi*.



Hình 4.5. Mô hình xem xét kỹ lưỡng

⁷ Petty, R. E., and Cacioppo, J. T. (1986). *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. New York: Springer-Verlag.

Việc quan điểm, thái độ cá nhân bị ảnh hưởng và thay đổi bởi hướng chính yếu hay bởi hướng ngoại vi phụ thuộc vào năng lực và động cơ của họ trong việc phân tích vấn đề. Chủ thể có năng lực phân tích tốt là những người có khả năng xử lý, phân tích thông tin một cách thấu đáo và có xu thế bị ảnh hưởng bởi các đặc điểm định tính của hiện tượng. Trong khi đó, những người hạn chế về khả năng phân tích thường bị ảnh hưởng bởi các tín hiệu ngoại vi. Lưu ý rằng, xu hướng phân tích là một đặc điểm có tính chất tình huống, chứ không phải là một tính cách cá nhân cố hữu. Ví dụ, một bác sĩ có thể sử dụng hướng chính yếu, kết hợp với chuyên môn của họ để chuẩn đoán và điều trị một loại bệnh nào đó. Tuy vậy, người bác sĩ này có thể dựa vào những tín hiệu ngoại vi từ bảng thông báo tự động trong xe để tìm hiểu điều gì đang xảy ra với xe hơi của họ. Chính vì sự đa dạng này, mô hình xem xét lý lưỡng được ứng dụng rộng rãi trong các chiến lược nhằm thay đổi thái độ của các cá nhân về một sản phẩm mới, ý tưởng mới hoặc thậm chí một thay đổi mới trong xã hội.

Lý thuyết răn đe tổng hợp (General Deterrence Theory). Hai nhà triết học duy lợi của thế kỷ XVIII là Cesare Beccaria và Jeremy Bentham đã xây dựng Lý thuyết răn đe tổng hợp để giải thích tình hình tội phạm và phương pháp làm giảm tội phạm. Lý thuyết răn đe tổng hợp xem xét lý do tại sao một số cá nhân tham gia vào các hành vi lệch chuẩn, chống đối xã hội hoặc phạm tội. Lý thuyết này cho rằng con người về cơ bản là duy lý (họ có lý trí khi thực hiện cả những hành vi phù hợp với chuẩn mực xã hội và cả hành vi lệch lạc) và con người được tự do lựa chọn hành vi lệch chuẩn dựa trên tính toán lợi ích hợp lý. Bởi theo lẽ tự nhiên, con người lựa chọn hành vi có lợi nhất, kể cả hành vi lệch chuẩn miễn là mang lại sự lợi ích hay khoái lạc cá nhân. Do đó, có thể được kiểm soát các hành vi đó bằng cách gia tăng mức độ trả giá cho hành vi lệch chuẩn ví dụ như tăng mức xử phạt (biện pháp đối phó) cũng như tăng khả năng bị phát hiện và trừng phạt đối với người vi phạm. Sự kịp thời, tính nghiêm khắc và chính xác của hình phạt là các yếu tố then chốt trong Lý thuyết răn đe tổng hợp.

Các nghiên cứu thực chứng cổ điển trong tội phạm học tìm kiếm nguyên nhân phổ biến của hành vi phạm tội như nghèo đói, thiếu giáo dục, các lý do tâm lý và sau đó đề xuất các chiến lược để giáo dục cải tạo tội phạm, chẳng hạn như tạo cho họ học nghề và điều trị y tế. Trong khi đó, Lý thuyết răn đe tổng hợp tập trung vào các quá trình quyết định phạm tội và các yếu tố tình huống ảnh hưởng đến quá trình lựa chọn hành vi phạm tội đó. Do đó, những yếu tố cá nhân của người phạm tội (như địa vị xã hội, sự sung túc và nhu cầu tiền bạc) và bối cảnh môi trường (như mức độ các mục tiêu được bảo vệ, năng lực của cảnh sát

khu vực, khả năng bị phát hiện và bắt giữ khi tội phạm xảy ra) đóng vai trò quan trọng trong quá trình ra quyết định phạm tội. Trọng tâm của Lý thuyết răn đe tổng hợp không phải là làm thế nào để giáo dục cải tạo tội phạm và ngăn chặn các hành vi tội phạm trong tương lai, mà là làm thế nào để làm giảm động cơ phạm tội (giảm mức độ hấp dẫn của hoạt động phạm tội) và qua đó ngăn ngừa tội phạm.

Để đạt được điều này, hàng loạt các biện pháp cụ thể được triển khai: thiết lập các “mục tiêu cứng” như các chốt cố định (deadbolts), xây dựng kỹ năng tự vệ, tăng tính răn đe của pháp luật như loại bỏ biện pháp tạm tha cho một số tội phạm, áp dụng “Luật bất quá tam” (“Three Strikes Law”)^(*) (buộc giam giữ đối với những đối tượng có 3 tiền án, ngay cả hành vi vi phạm thuộc loại ít nghiêm trọng và trong khung được miễn chấp hành hình phạt tù), áp dụng án tử hình, tăng nhận thức về khả năng bị phát hiện, bắt giữ bằng cách bố trí phương tiện camera dọc các tuyến phố, sử dụng lực lượng đặc nhiệm chống ma túy hay tội phạm có tổ chức, tăng cường tuần tra cảnh sát, thực hiện các chương trình giáo dục như thông báo rộng rãi khẩu hiệu “Mọi hành vi phạm tội sẽ bị trừng phạt”. Lý thuyết này có ý nghĩa không chỉ cho tội phạm truyền thống, mà còn cho tội phạm cổ cồn trắng hiện đại, như mua bán thông tin nội bộ trong kinh doanh chứng khoán, vi phạm bản quyền phần mềm, chia sẻ bất hợp pháp tác phẩm âm nhạc.

^(*) “Three Strike Law” là một đạo luật rất nghiêm khắc của một số bang Hoa Kỳ, tạm dịch là “Luật bất quá tam” có nghĩa là “Luật ba lần phạm tội hình sự là bị xử tù chung thân” (N.D).

Chương 5

THIẾT KẾ NGHIÊN CỨU

Thiết kế nghiên cứu là một kế hoạch toàn diện, chỉ rõ cách thức thu thập dữ liệu trong một dự án nghiên cứu thực nghiệm. Thiết kế nghiên cứu là một “lộ trình chi tiết” cho các nghiên cứu thực nghiệm nhằm trả lời các câu hỏi nghiên cứu cụ thể hoặc kiểm nghiệm các giả thuyết cụ thể và phải ghi rõ ít nhất ba quá trình: (1) quá trình thu thập dữ liệu, (2) quá trình thuyết minh công cụ sử dụng và (3) quá trình lấy mẫu. Quá trình thuyết minh công cụ sử dụng và quá trình lấy mẫu sẽ được mô tả trong hai chương tiếp theo, quá trình thu thập dữ liệu (đôi khi còn được gọi là quá trình “thiết kế nghiên cứu”) được giới thiệu trong chương này và được mô tả chi tiết hơn trong các Chương 9 - 12.

Nhìn chung, các phương pháp thu thập dữ liệu có thể được nhóm lại thành hai loại: thực chứng và diễn giải. **Phương pháp thực chứng** (positivist method), chẳng hạn như các thực nghiệm trong phòng thí nghiệm và nghiên cứu khảo sát, nhằm mục đích kiểm nghiệm lý thuyết (hay giả thuyết), trong khi **phương pháp diễn giải** (interpretive method), chẳng hạn như nghiên cứu hành động và dân tộc học, nhằm mục đích hình thành lý thuyết. Phương pháp thực chứng sử dụng cách tiếp cận suy diễn để nghiên cứu, bắt đầu với một lý thuyết và thử nghiệm các luận điểm lý thuyết bằng cách sử dụng các dữ liệu thực nghiệm. Ngược lại, phương pháp diễn giải sử dụng cách tiếp cận quy nạp bắt đầu với dữ liệu và cố gắng để tạo ra một lý thuyết về hiện tượng quan tâm từ các dữ liệu đã thu thập.

Thông thường, những phương pháp này không hoàn toàn tương đương với nghiên cứu định lượng và định tính. Phương pháp định lượng và định tính đề cập đến các loại dữ liệu được thu thập (dữ liệu định lượng liên quan đến chỉ số, số liệu; trong khi các dữ liệu định tính bao gồm các cuộc phỏng vấn, quan sát,...) và được phân tích (ví dụ, các kỹ thuật định lượng như hồi quy hoặc các kỹ thuật định tính như mã hoá). Nghiên cứu thực chứng sử dụng dữ liệu chủ yếu là định lượng, nhưng cũng có thể sử dụng dữ liệu định tính. Nghiên cứu diễn giải chủ yếu dựa vào các dữ liệu định tính, nhưng đôi khi có thể được hưởng lợi từ các dữ liệu định lượng. Đôi khi, sử dụng kết hợp các dữ liệu định tính và định lượng có thể giúp tạo ra cái nhìn sâu sắc đối với một hiện tượng xã hội phức tạp mà nếu chỉ dùng một trong hai loại dữ liệu trên thì không thể có được. Do đó, các thiết kế hỗn hợp kết hợp các dữ liệu định tính và định lượng thường là rất hấp dẫn.

Các thuộc tính cơ bản của một thiết kế nghiên cứu

Chất lượng của các thiết kế nghiên cứu có thể được xác định dựa vào bốn yếu tố chính: giá trị nội tại, giá trị ngoại tại, giá trị cấu trúc và giá trị kết luận thống kê.

Giá trị nội tại (internal validity), cũng được gọi là *giá trị nhân quả*, xem xét liệu sự thay đổi của một biến phụ thuộc được quan sát có thực sự được tạo ra bởi một sự thay đổi của biến độc lập tương ứng hay đó là kết quả gây ra bởi các biến làm nhiễu xuất hiện trong bối cảnh nghiên cứu. Giá trị nhân quả đòi hỏi thoả mãn ba điều kiện: (1) sự đồng biến của nguyên nhân và kết quả (ví dụ, nếu nguyên nhân xảy ra, thì hậu quả cũng sẽ xảy ra; nếu nguyên nhân không xảy ra, hậu quả không xảy ra), (2) thời gian ưu tiên: nguyên nhân phải đi trước kết quả về mặt thời gian, (3) không có lời giải thích chính đáng khác (nghĩa là không có tương quan giả).

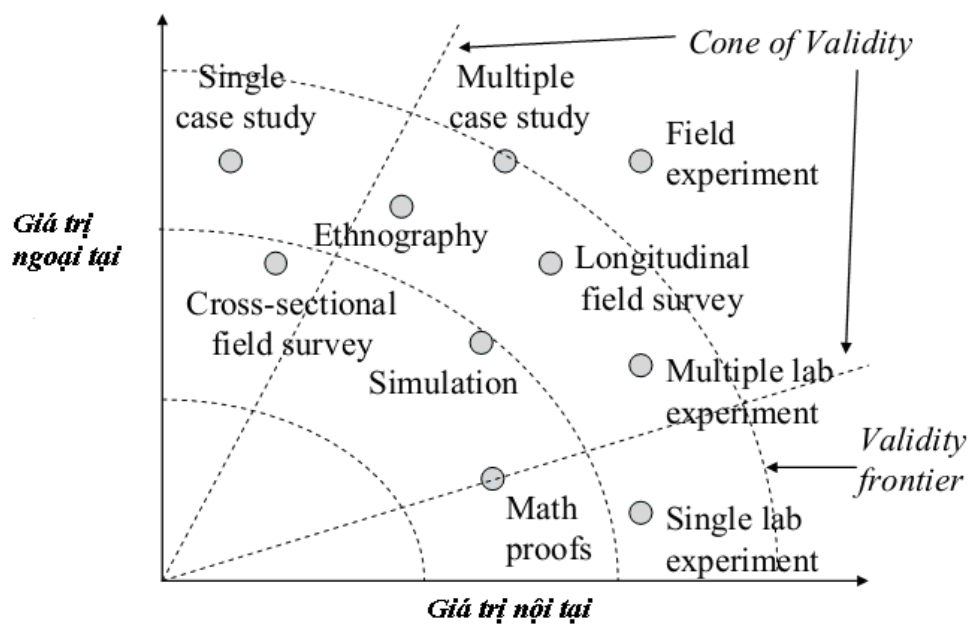
Một số thiết kế nghiên cứu, chẳng hạn như các thử nghiệm trong phòng thí nghiệm có giá trị nhân quả cao nhờ khả năng xử lý, điều chỉnh các biến độc lập (nguyên nhân), sau đó đo lường chính xác kết quả quan sát (biến phụ thuộc), cùng với việc kiểm soát những tác động của các biến không liên quan (biến nhiễu). Các thiết kế khác, chẳng hạn như khảo sát thực địa, thường không đạt được giá trị nội tại cao bởi vì trước hết chúng không có khả năng điều chỉnh các biến độc lập (nguyên nhân). Ngoài ra, trong khảo sát thực địa, khi nguyên nhân và kết quả được đánh giá tại cùng một điểm và như vậy có thể xảy ra tình huống trước khi khảo sát, yếu tố kết quả (biến phụ thuộc) đã có tác động, ảnh hưởng tới yếu tố nguyên nhân hoặc đã bị tác động bởi các yếu tố khác.

Việc có giá trị nội tại (giá trị nhân quả) cao hơn so với các thiết kế khác không có nghĩa là thực nghiệm trong phòng thí nghiệm tránh được các mối đe dọa đối với giá trị nhân quả. Trên thực tế có hàng loạt các nguy cơ là giảm giá trị nhân quả trong nghiên cứu thực nghiệm. Đó là những ảnh hưởng bởi lịch sử, quá trình thử nghiệm, thiết bị đo đạc, hồi quy,... Tất cả sẽ được thảo luận sau trong chương về thiết kế thực nghiệm.

Giá trị ngoại tại (external validity)^(*) hay giá trị mở rộng, đề cập tới khả năng các kết quả nghiên cứu có thể được suy rộng cho các quần thể khác (giá trị quần thể) hay cho các trường hợp khác: con người khác, tổ chức khác, bối cảnh khác và thời gian khác (giá trị sinh thái). Ví dụ, liệu rằng các kết quả rút ra từ một mẫu nghiên cứu của các công ty tài chính ở Hoa Kỳ có thể được áp dụng cho các công ty tài chính khác trên thế giới (giá trị

(*) Nguyễn Văn Tuấn dịch “internal validity” là “hợp lý nội tại”; “external validity” là “hợp lý ngoại tại”. Xem: “Tiêu chuẩn chất lượng giáo dục đại học”, ngày 02/02/2008. Nguồn: tiasang.com.vn/Default.aspx?tabid=62&News=1547&CategoryID=6

quần thể) hoặc cho các công ty khác tại Hoa Kỳ (giá trị sinh thái)? Nghiên cứu khảo sát với dữ liệu được thu thập từ nhiều nguồn khác nhau như các cá nhân, tổ chức hoặc các đơn vị phân tích khác, thường có tính khái quát hóa cao hơn so với các thử nghiệm trong phòng thí nghiệm với các tác động chủ quan của nhà nghiên cứu và sự kiểm soát chặt chẽ các biến không liên quan. Điều này sẽ tạo ra các kết quả ít khái quát hơn bởi vì trong thực tế nhà nghiên cứu không thể kiểm soát được các biến không liên quan này. Sự biến động về giá trị nội tại và giá trị ngoại tại của các thiết kế nghiên cứu được trình bày trong Hình 5.1.



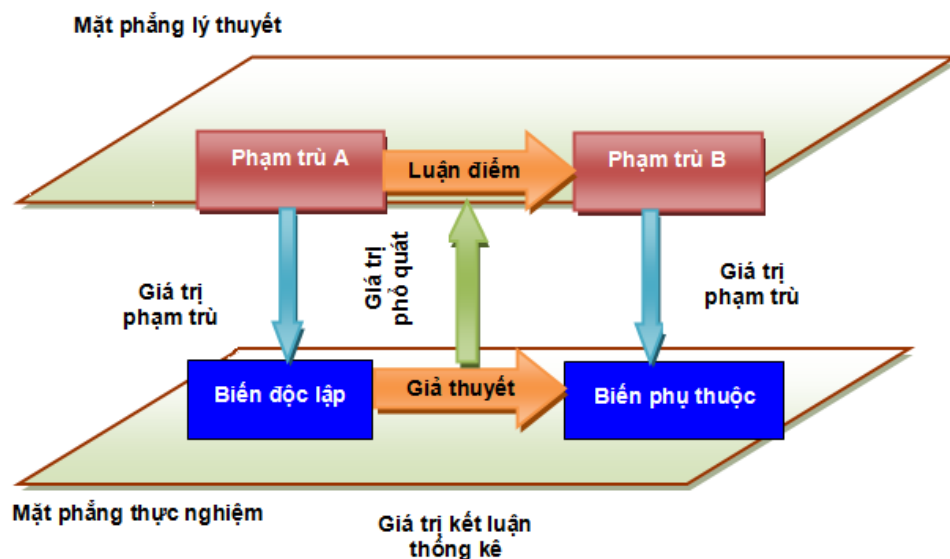
Hình 5.1. Giá trị nội tại và giá trị ngoại tại

Một số nhà nghiên cứu khẳng định rằng có sự trao đổi giữa giá trị nội tại và giá trị ngoại tại: nếu muốn giá trị ngoại tại cao thì phải chấp nhận giá trị nội tại thấp và ngược lại. Nhưng điều này không phải luôn luôn đúng. Các thiết kế nghiên cứu như thực nghiệm ngoài thực tế, khảo sát thực tế dài hơi và các nghiên cứu đa trường hợp mang đến giá trị cao cả về nội tại lẫn ngoại tại. Cá nhân tôi thích các thiết kế nghiên cứu có mức độ hợp lý của giá trị nội tại và ngoại tại - đó là những nghiên cứu nằm trong hình nón như ở Hình 5.1. Nhưng điều này không có nghĩa là các thiết kế bên ngoài hình nón này là kém giá trị và ít hữu ích. Các nhà nghiên cứu lựa chọn các thiết kế với mong muốn phù hợp nhất với sở thích, năng lực cá nhân; mức độ giá trị nội tại và ngoại tại mà họ mong muốn.

Giá trị phạm trù (construct validity) xem xét có đúng là thang đánh giá được sử dụng để đo lường, đánh giá các phạm trù lý thuyết cần nghiên cứu, chứ không phải là các phạm trù khác. Nhiều phạm trù được sử dụng trong nghiên cứu khoa học xã hội như sự

đồng cảm, bảo thủ và học tập cộng đồng là rất khó để định nghĩa, càng khó để đo lường. Ví dụ, giá trị phạm trù cần đảm bảo chắc chắn một thang đánh giá thực sự để đo lường “sự đồng cảm”, chứ không phải là “lòng từ bi”. Điều này có thể khó khăn khi nhiều phạm trù có ý nghĩa tương tự. Giá trị phạm trù được đánh giá trong nghiên cứu thực chứng dựa trên phương pháp phân tích tương quan hoặc phân tích giai thừa và sẽ được mô tả trong chương kế tiếp.

Giá trị kết luận thống kê (statistical conclusion validity) đánh giá mức độ xác thực của các kết luận có được bằng việc sử dụng các kỹ thuật thống kê. Ví dụ, tiêu chuẩn này sẽ kiểm tra liệu rằng phương pháp thống kê được sử dụng để kiểm nghiệm các giả thuyết có phù hợp, liệu rằng các biến được sử dụng có đáp ứng được các điều kiện để áp dụng phương pháp thống kê đó (chẳng hạn như kích cỡ tối thiểu của mẫu, yêu cầu về phân bố mẫu). Bởi các thiết kế nghiên cứu diễn giải không sử dụng các kỹ thuật kiểm định thống kê, nên giá trị kết luận thống kê không áp dụng cho phân tích này. Các giá trị khác nhau và vị trí chúng dưới góc độ lý thuyết / thực nghiệm được minh họa trong Hình 5.2.



Hình 5.2. Kiểu giá trị khác nhau trong nghiên cứu khoa học

Cải thiện giá trị nội tại và giá trị ngoại tại

Những thiết kế nghiên cứu tốt nhất là những thiết kế có khả năng đảm bảo giá trị cao cả về nội tại lẫn ngoại tại. Những thiết kế như vậy sẽ hạn chế được các quan hệ không xác thực, nâng cao niềm tin trong việc kiểm nghiệm giả thuyết và đảm bảo các kết quả có được

từ một mẫu nhỏ có thể đại diện cho cả một quần thể lớn. Cần phải có được những đối chứng cần thiết trong thiết kế nghiên cứu để nâng cao giá trị nội tại (nhân quả). Có năm cách để có được sự đối chứng là: (1) điều chỉnh, (2) loại bỏ, (3) bao hàm, (4) đối chứng thống kê và (5) ngẫu nhiên hóa.

Trong kỹ thuật **điều chỉnh** (manipulation), nhà nghiên cứu sử dụng và điều chỉnh các thử nghiệm đối với các biến độc lập ở các mức độ khác nhau (gọi là các mức độ “thử nghiệm”); sau đó so sánh ảnh hưởng của các thử nghiệm này với một nhóm đối chứng (nhóm này không tiếp nhận các tác động thử nghiệm). Ví dụ về thử nghiệm có thể là một loại thuốc mới, một liều lượng mới, một phương pháp giảng dạy mới,... Kỹ thuật điều chỉnh sẽ phát huy tác dụng trong các thiết kế thực nghiệm (experimental) và ngụy thực nghiệm (quasi-experimental). Trong khi đó ở các thiết kế phi thực nghiệm (non-experimental) như khảo sát chẳng hạn, kỹ thuật này khó có thể áp dụng. Nếu đối tượng nghiên cứu không thể phân biệt được các mức độ khác nhau của quá trình điều tiết, thì phản ứng của họ cho các mức độ này có thể sẽ không khác nhau. Như vậy kỹ thuật điều chỉnh sẽ thất bại.

Kỹ thuật loại trừ (elimination) nhằm loại bỏ những biến không liên quan bằng cách liên tục kiểm tra tính liên quan của các biến, chẳng hạn như bằng cách chỉ tập trung nghiên cứu một giới tính hoặc một nhóm khách thể có cùng điều kiện kinh tế - xã hội. Trong **kỹ thuật bao hàm** (inclusion), các biến không liên quan được sử dụng, sau đó mức độ tác động của chúng đối với các biến độc lập sẽ được tính toán cụ thể. Phương pháp này giúp tăng giá trị phổ quát, nhưng tất nhiên đòi hỏi số lượng mẫu phải lớn hơn nhiều đủ để kiểm tra nhiều biến khác nhau. Trong phương pháp **đối chứng thống kê** (statistical control), các biến không liên quan được sử dụng, đo lường như những đồng biến trong quá trình thử nghiệm.

Cuối cùng, trong **kỹ thuật ngẫu nhiên hóa** (randomization), tác động của các biến không liên quan sẽ được loại trừ thông qua quá trình lấy mẫu ngẫu nhiên. Có hai cách ngẫu nhiên hoá (1) lựa chọn ngẫu nhiên tức là mẫu được lấy ngẫu nhiên từ quần thể, và (2) phân bổ ngẫu nhiên là đối tượng nghiên cứu được lựa chọn không ngẫu nhiên nhưng lại được phân bổ ngẫu nhiên vào các nhóm thử nghiệm. Ngẫu nhiên hoá cũng sẽ cải thiện giá trị phổ quát, cho phép kết luận từ mẫu có thể được khái quát hoá cho toàn bộ quần thể. Chú ý rằng, nếu lựa chọn ngẫu nhiên không thực hiện được vì lý do hạn chế về nguồn lực và việc tiếp cận đối tượng nghiên cứu thì phân bổ ngẫu nhiên là bắt buộc. Bên cạnh đó, việc khái

quát hoá cũng sẽ bị hạn chế khi quần thể gồm nhiều khía cạnh đa dạng phức tạp, trong khi đó chúng ta chỉ có thể kiểm soát được một vài khía cạnh này mà thôi.

Các thiết kế nghiên cứu phổ biến

Như đã đề cập phân trước, tùy thuộc vào mục đích nghiên cứu, thiết kế nghiên cứu có thể được phân loại thành hai loại - thực chứng và diễn giải. Thiết kế thực chứng nhằm mục đích thử nghiệm lý thuyết, trong khi các thiết kế diễn giải nhằm xây dựng lý thuyết. Thiết kế thực chứng tìm kiếm mô hình tổng quát dựa trên một cái nhìn khách quan về thực tế, trong khi các thiết kế diễn giải tìm kiếm giải thích chủ quan về các hiện tượng xã hội từ quan điểm của các đối tượng liên quan.

Một số ví dụ phổ biến của các thiết kế thực chứng bao gồm thử nghiệm trong môi trường nhân tạo, thử nghiệm ngoài thực địa, khảo sát thực địa, phân tích dữ liệu thứ cấp và nghiên cứu trường hợp; trong khi ví dụ về thiết kế diễn giải bao gồm nghiên cứu trường hợp, hiện tượng học và dân tộc học. Lưu ý rằng nghiên cứu trường hợp có thể được sử dụng để xây dựng lý thuyết hoặc thử nghiệm lý thuyết. Không phải tất cả các phương pháp này phù hợp cho tất cả các loại nghiên cứu khoa học. Một số kỹ thuật như hội thảo nhóm nhỏ là thích hợp nhất cho các nghiên cứu thăm dò, những kỹ thuật khác như nghiên cứu dân tộc học phát huy tác dụng tốt nhất cho nghiên cứu mô tả, hay các thử nghiệm trong môi trường nhân tạo là nơi lý tưởng cho nghiên cứu giải thích. Sau đây là mô tả ngắn gọn về một số những thiết kế này. Các chương 9 - 12 sẽ cung cấp chi tiết hơn.

Nghiên cứu thực nghiệm (experimental studies)^(*) là những nghiên cứu nhằm mục đích kiểm tra mối quan hệ nhân - quả (các giả thuyết) trong điều kiện, môi trường nghiên cứu được kiểm soát chặt chẽ bằng cách tách biệt các yếu tố nguyên nhân với kết quả về mặt thời gian, áp đặt yếu tố nguyên nhân đối với một nhóm đối tượng (nhóm thực nghiệm) mà không áp đặt đối với một nhóm khác (nhóm đối chứng). Sau đó quan sát sự thay đổi của các kết quả giữa các đối tượng trong hai nhóm này. Ví dụ, nếu chúng ta thiết kế một thực nghiệm trong phòng thí nghiệm để kiểm tra tính hiệu quả của một loại thuốc mới trong điều trị một căn bệnh nào đó, có thể lấy một mẫu ngẫu nhiên trong số những người bị ảnh hưởng với căn bệnh đó. Sau đó phân họ một cách ngẫu nhiên thành hai nhóm (nhóm thực nghiệm và nhóm đối chứng), phát thuốc cho các đối tượng trong nhóm điều trị (nhóm thực nghiệm), nhưng chỉ đưa ra viên trấn an (viên thuốc không có giá trị chữa bệnh) cho nhóm khác. Các thiết kế phức tạp hơn có thể bao gồm nhiều nhóm thực nghiệm, chẳng hạn

^(*)“Experimental studies” có khi được dùng với nghĩa là nghiên cứu thử nghiệm hay nghiên cứu thí nghiệm

như các nhóm được phát thuốc với liều lượng cao thấp khác nhau; hoặc sử dụng nhiều tác động, chẳng hạn như kết hợp dùng thuốc với việc can thiệp vào chế độ ăn uống.

Trong **thiết kế thực nghiệm đích thực** (true experimental design), các đối tượng này phải được phân ngẫu nhiên giữa các nhóm. Nếu việc phân bổ ngẫu nhiên không được thực hiện thì thiết kế đó sẽ trở thành **giả thực nghiệm** (quasi-experiment). Các thực nghiệm có thể được thực hiện trong một môi trường nhân tạo hoặc phòng thí nghiệm như của một trường đại học (thực nghiệm nhân tạo - laboratory experiments) hoặc ngoài thực địa như trong một tổ chức nơi các hiện tượng quan tâm đang diễn ra trên thực tế (thực nghiệm thực địa - field experiments). Thực nghiệm trong phòng thí nghiệm cho phép các nhà nghiên cứu có thể dễ dàng cô lập các biến quan tâm và kiểm soát các biến không liên quan, điều này khó có thể có được trong các thực nghiệm thực địa. Do đó, kết luận rút ra từ các thực nghiệm nhân tạo thường có giá trị nhân quả (giá trị nội tại) cao hơn, nhưng các thực nghiệm thực địa lại thường có giá trị ngoại tại cao hơn.

Dữ liệu thực nghiệm được phân tích bằng cách sử dụng các kỹ thuật thống kê định lượng. Điểm mạnh nhất của thiết kế thực nghiệm là giá trị nhân quả cao do khả năng cô lập, kiểm soát và tập trung kiểm tra một số lượng nhỏ các biến, trong khi điểm yếu cơ bản của nó là sự hạn chế về khả năng khái quát hóa vì trên thực tế cuộc sống thường phức tạp hơn (ví dụ như có sự xuất hiện của nhiều biến không liên quan) so với các thiết lập trong phòng thí nghiệm. Hơn nữa, nếu nghiên cứu không nhận diện và kiểm soát các biến không liên quan thì giá trị nhân quả có thể bị giảm hoặc mang đến những tương quan giả.

Khảo sát thực địa (field survey) là những thiết kế phi thực nghiệm, không kiểm soát hay tác động các biến độc lập mà đánh giá các biến này và kiểm tra kết quả bằng cách sử dụng các phương pháp thống kê. Khảo sát thực địa giúp nắm bắt được bức tranh tổng thể của thực tế, các quan điểm, tình huống từ một mẫu ngẫu nhiên của các đối tượng trong các môi trường nghiên cứu thông qua một bảng câu hỏi khảo sát hoặc ít thường xuyên hơn, thông qua một cuộc phỏng vấn cấu trúc (structured interview). Trong các cuộc **khảo sát thực địa ngang** (cross-sectional field survey), các biến độc lập và biến phụ thuộc được đánh giá tại cùng một điểm (ví dụ, bằng cách sử dụng một bảng câu hỏi duy nhất), trong khi đó đối với **khảo sát thực địa dọc** (còn gọi là khảo sát thực địa dài hơi - longitudinal field survey), các biến phụ thuộc được đánh giá sau các biến độc lập. Thế mạnh của khảo sát thực địa là giá trị phổ quát của chúng (bởi vì dữ liệu được thu thập trong thực tế), có khả năng nắm bắt và kiểm soát một số lượng lớn các biến và khả năng nghiên cứu một vấn đề từ nhiều bối cảnh hoặc sử dụng các nhiều lý thuyết khác nhau. Tuy nhiên, do tính đa

dạng, phức tạp của môi trường thực địa, giá trị nhân quả (mối quan hệ nhân - quả) rất khó được suy ra và các cuộc khảo sát có thể gặp phải những thành kiến của đối tượng khảo sát (ví dụ như đối tượng có thể cung cấp những câu trả lời theo xu thế xã hội mà không phải là phản ứng thật sự của họ). Điều này làm giảm giá trị nhân quả.

Phân tích dữ liệu thứ cấp (secondary data analysis) là việc phân tích các dữ liệu được người khác thu thập và xử lý trước đây. Những dữ liệu này có thể bao gồm dữ liệu từ các cơ quan chính phủ như các số liệu thống kê việc làm từ các Cơ quan Dịch vụ Lao động, thống kê phát triển quốc gia từ các Chương trình Phát triển Liên Hợp Quốc, dữ liệu do các nhà nghiên cứu khác thu thập (thường được sử dụng trong các nghiên cứu phân tích tổng hợp) hoặc dữ liệu của bên thứ ba được công bố công khai, chẳng hạn như dữ liệu tài chính từ thị trường chứng khoán hoặc dữ liệu đấu giá từ eBay. Điều này là trái ngược với hầu hết các thiết kế nghiên cứu khác khi mà thu thập dữ liệu là một phần công việc của nhà nghiên cứu.

Phân tích dữ liệu thứ cấp có thể là một phương tiện nghiên cứu hiệu quả khi mà việc thu thập dữ liệu gốc quá tốn kém hoặc không khả thi và dữ liệu thứ cấp có sẵn, đủ về số lượng và chất lượng để trả lời các câu hỏi nghiên cứu. Những hạn chế của thiết kế này là một số nguồn dữ liệu có thể không được thu thập một cách có hệ thống, khoa học và do đó không phù hợp với nghiên cứu khoa học. Rõ ràng nếu dữ liệu được thu thập cho mục đích của người khác thì chúng khó có thể giải quyết triệt để các câu hỏi nghiên cứu cho mục đích của bạn.

Nghiên cứu trường hợp (Case study) là việc điều tra sâu, trong một khoảng thời gian dài về một hoặc nhiều vấn đề khoa học được phát hiện trong thực tế cuộc sống. Dữ liệu có thể được thu thập bằng cách sử dụng kết hợp các phương pháp như phỏng vấn, quan sát cá nhân và phân tích tài liệu. Về bản chất, nghiên cứu trường hợp có thể là thực chứng (dùng để kiểm nghiệm giả thuyết) hoặc diễn giải (dùng để xây dựng lý thuyết). Điểm mạnh của phương pháp nghiên cứu này là khả năng khám phá một loạt các yếu tố xã hội, văn hóa và chính trị có khả năng liên quan đến hiện tượng cần nghiên cứu chưa được khám phá trước đó. Nghiên cứu trường hợp chứa đựng nhiều các phân tích định tính, phụ thuộc nhiều vào bối cảnh chứa đựng trường hợp nghiên cứu. Tuy nhiên, việc giải thích các kết quả phụ thuộc nhiều vào khả năng quan sát và tích hợp của nhà nghiên cứu. Việc thiếu so sánh đối chứng có thể tạo ra những khó khăn trong thiết lập mối quan hệ nhân quả. Đồng thời, những phát hiện từ trường hợp đơn lẻ có thể không dễ dàng để khái quát hóa cho các trường hợp khác. Khái quát hóa có thể được cải thiện bằng cách tái tạo, so sánh và

kết hợp với các trường hợp khác trong trường hợp một *thiết kế nhiều trường hợp* (*multi case design*).

Thảo luận nhóm nhỏ (còn gọi thảo luận nhóm tâm điểm - Focus group research) là việc tập hợp một nhóm nhỏ người tham gia (thường là 6 đến 10 người) tại một địa điểm và tổ chức thảo luận về một hiện tượng quan tâm trong khoảng thời gian 1,5 đến 2 giờ. Thảo luận này được chủ trì và điều hành bởi một người có chuyên môn. Người điều hành thiết lập chương trình thảo luận và đặt ra một tập hợp các câu hỏi ban đầu cho người tham gia. Quá trình thảo luận, người này phải đảm bảo chắc chắn rằng những ý tưởng và kinh nghiệm của tất cả những người tham gia phải được phát biểu. Sau đó phải cố gắng giải quyết vấn đề nghiên cứu một cách toàn diện dựa trên việc tổng hợp các ý kiến và kinh nghiệm của người tham gia. Trong thảo luận nhóm nhỏ, giá trị nhân quả không thể được thiết lập do thiếu sự đối chứng và phát hiện không thể được khái quát hóa cho các đối tượng khác vì kích thước mẫu rất nhỏ. Do đó, thảo luận nhóm nhỏ thường không được sử dụng trong nghiên cứu giải thích hoặc mô tả mà thiết kế này phù hợp hơn cho các nghiên cứu thăm dò.

Nghiên cứu hành động (Action research) giả định rằng cách tốt nhất để nhận thức các hiện tượng xã hội phức tạp là trực tiếp can thiệp, tác động vào những hiện tượng đó và sau đó quan sát kết quả của quá trình tác động. Trong phương pháp này, các nhà nghiên cứu thường là một nhà tư vấn hoặc là một thành viên sẽ tham dự vào một bối cảnh xã hội nhất định, ví dụ như họ tham gia vào một tổ chức, đưa ra những đề xuất để cải thiện hoạt động của tổ chức đó. Việc lựa chọn hành vi của nhà nghiên cứu phải dựa trên những cơ sở lý thuyết; phải giải thích được lý do tại sao và làm thế nào hành vi đó có thể gây ra các thay đổi theo mong muốn. Nhà nghiên cứu sau đó quan sát các kết quả tác động bởi hành vi của họ, sửa đổi hành vi can thiệp khi cần thiết, đồng thời phân tích để đưa ra những hiểu biết lý thuyết về các vấn đề nghiên cứu và các biện pháp can thiệp đã thực hiện.

Lý thuyết ban đầu có giá trị nếu hành vi được lựa chọn giải quyết thành công vấn đề nghiên cứu. Giải quyết được vấn đề nghiên cứu tương tự, đồng đại là các đặc tính căn bản để phân biệt nghiên cứu hành động với tất cả các phương pháp nghiên cứu khác. Do đó, nghiên cứu hành động là một phương pháp được đánh giá cao, làm cầu nối giữa nghiên cứu và thực hành. Phương pháp này cũng thích hợp cho việc nghiên cứu các vấn đề xã hội đặc thù không tìm thấy bên ngoài bối cảnh của chúng. Tuy vậy, phương pháp này cũng có thể mang đến những thiên vị chủ quan của nhà nghiên cứu và việc khái quát hóa những kết quả thường bị giới hạn trong bối cảnh nghiên cứu được tiến hành.

Nghiên cứu dân tộc học (Ethnography) là một thiết kế nghiên cứu diễn giải lấy cảm hứng từ nhân chủng học. Nghiên cứu dân tộc học nhấn mạnh rằng các hiện tượng nghiên cứu phải được nghiên cứu trong bối cảnh nền văn hóa của chúng. Nhà nghiên cứu cần phải đắm mình trong một nền văn hóa nhất định trong thời gian dài (8 tháng đến 2 năm) và trong thời gian đó họ tham gia, quan sát, cảm nhận và ghi lại cuộc sống hàng ngày; sau cùng hình thành lý thuyết về sự vận động của nền văn hóa đó. Dữ liệu được thu thập chủ yếu thông qua các kỹ thuật quan sát, quá trình tương tác chính thức và không chính thức với những con người trong nền văn hóa đó và do vậy, những ghi chú cá nhân có ý nghĩa quan trọng trong phân tích dữ liệu.

Các nhà nghiên cứu phải tường thuật rất chi tiết kinh nghiệm của mình để người đọc có thể có những trải nghiệm giống như nhà nghiên cứu về nền văn hóa mà không nhất thiết phải có mặt ở đó. Những lợi thế của cách tiếp cận này là sự nhạy cảm trong bối cảnh, sự hiểu biết phong phú, có sắc thái và hạn chế thấp nhất những thiên vị hay định kiến của khách thể nghiên cứu. Tuy nhiên, đây cũng là một cách tiếp cận cực kỳ tốn thời gian, công sức và kinh phí, những phát hiện chỉ đặc trưng cho một nền văn hóa nhất định và ít khái quát hóa cho các nền văn hóa khác.

Lựa chọn các thiết kế nghiên cứu

Trong số rất nhiều trên thiết kế nghiên cứu nêu trên, nhà nghiên cứu nên chọn thiết kế nào cho nghiên cứu của họ? Nhìn chung, các nhà nghiên cứu có xu hướng chọn những thiết kế mà họ thấy thuận lợi nhất và tự tin nhất để thực hiện. Nhưng về nguyên tắc, việc lựa chọn nên căn cứ vào bản chất của hiện tượng cần nghiên cứu chứ không phải là sự tiện lợi của nhà nghiên cứu. Trong giai đoạn nghiên cứu ban đầu, khi các vấn đề nghiên cứu không rõ ràng, nhà nghiên cứu muốn tìm hiểu kỹ tính chất và mức độ của một vấn đề nghiên cứu nhất định, thảo luận nhóm nhỏ (đối với đơn vị phân tích là cá nhân) hoặc nghiên cứu trường hợp (đối với đơn vị phân tích tổ chức) là một chiến lược lý tưởng cho các nghiên cứu thăm dò.

Khi muốn đào sâu hơn nữa vào lĩnh vực nghiên cứu, nhưng thấy không có lý thuyết thuyết phục để giải thích các hiện tượng quan tâm và muốn xây dựng một lý thuyết để khóa lấp các khoảng trống lý luận trong lĩnh vực đó, các thiết kế diễn giải như nghiên cứu trường hợp hoặc nghiên cứu dân tộc học có thể rất hữu ích. Nếu các lý thuyết hiện có đang tồn tại những mâu thuẫn và nhà nghiên cứu muốn kiểm nghiệm các lý thuyết khác nhau hoặc tích hợp chúng thành một lý thuyết lớn hơn, các thiết kế thực chứng như thực nghiệm,

nghiên cứu khảo sát, phân tích dữ liệu thứ cấp lại phù hợp hơn.

Cho dù thiết kế nghiên cứu cụ thể nào được chọn, các nhà nghiên cứu cần phải cố gắng để thu thập cả dữ liệu định tính và định lượng bằng cách sử dụng kết hợp các kỹ thuật như bảng câu hỏi, phỏng vấn, quan sát và dữ liệu thứ cấp. Ví dụ, ngay cả trong một bảng câu hỏi cố định để thu thập dữ liệu định lượng, nhà nghiên cứu có thể để lại một số chỗ cho một vài câu hỏi mở để thu thập dữ liệu định tính, điều này có thể thu nhận những kết quả bất ngờ không thể có được nếu chỉ sử dụng các câu hỏi đã cấu trúc sẵn trước đó. Tương tự như vậy, trong khi nghiên cứu trường hợp sử dụng chủ yếu là phỏng vấn trực tiếp để thu thập dữ liệu định tính thì cũng không nên bỏ qua tiềm năng và giá trị của việc thu thập dữ liệu định lượng.

Ví dụ, trong nghiên cứu trường hợp về quá trình ra quyết định của một tổ chức, khi phỏng vấn (thu thập dữ liệu định tính) nhà nghiên cứu có thể ghi lại các con số như mất bao nhiêu tháng để đưa ra quyết định, bao nhiêu người đã tham gia trong quá trình ra quyết định và bao nhiêu quyết định đã bị thay thế. Việc làm này có thể cung cấp những hiểu biết có giá trị mà không dễ có trong các câu trả lời tường thuật của người được hỏi. Cho dù thiết kế nghiên cứu cụ thể nào được sử dụng, nhà nghiên cứu cần phải hướng đến mục tiêu thu thập dữ liệu nhiều và đa dạng nhất có thể, có như vậy mới có thể tạo ra những cái nhìn sâu sắc nhất về hiện tượng nghiên cứu.

Chương 6

PHƯƠNG THỨC ĐÁNH GIÁ

Luận điểm lý thuyết chứa đựng mối quan hệ giữa các phạm trù trừu tượng. Kiểm nghiệm lý thuyết (kiểm tra các luận điểm) đòi hỏi phải đo lường, đánh giá các phạm trù đó một cách đầy đủ, chính xác bằng các phương pháp khoa học. Trong nghiên cứu khoa học xã hội, một số phạm trù, chẳng hạn như độ tuổi hay trọng lượng của một người, quy mô của một tổ chức, có thể dễ dàng đo lường, trong khi các phạm trù khác, chẳng hạn như thành kiến, sáng tạo hay sự xa lánh thì việc đo lường, đánh giá có thể khó khăn hơn nhiều. Trong chương này, chúng ta sẽ tìm hiểu các quy trình liên quan đến việc nhận thức (khái niệm hóa) và thực hành (thao tác hóa) để tạo ra các biện pháp đánh giá các phạm trù.

Khái niệm hóa

Khái niệm hóa (conceptualization) là quá trình nhận thức, trong đó các phạm trù (hay khái niệm) và các thành tố cấu thành của chúng còn mơ hồ, chưa chính xác được định nghĩa bằng những thuật ngữ cụ thể, rõ ràng và chính xác. Ví dụ, chúng ta thường sử dụng từ “định kiến” (prejudice) và thuật ngữ này gợi lên những hình dung khác nhau trong nhận thức của chúng ta, tuy nhiên, chúng ta có thể cảm thấy hết sức khó khăn để xác định chính xác thuật ngữ này có nghĩa là gì. Nếu ai đó nói những điều xấu về các chủng tộc khác, liệu đó có phải là định kiến về chủng tộc? Nếu phụ nữ được trả ít tiền hơn nam giới khi họ làm cùng một công việc, đó có phải là định kiến về giới? Nếu các con chiên tin rằng người không có niềm tin vào Chúa sẽ bị thiêu đốt trong địa ngục. Vậy đó có phải là định kiến tôn giáo? Có các loại định kiến khác nhau và chúng là những gì? Có những mức độ khác nhau của định kiến, chẳng hạn như sâu sắc hay chỉ một chút đỉnh? Trả lời tất cả những câu hỏi này là chìa khóa để đánh giá, đo lường một cách chính xác phạm trù “định kiến”. Quá trình nhận thức về những gì sẽ được bao hàm và những gì được loại trừ trong khái niệm “định kiến” chính là quá trình khái niệm hóa.

Quá trình hình thành khái niệm là điều kiện tiên quyết trong khoa học xã hội bởi vì trong thực tiễn có rất nhiều phạm trù thiếu chính xác, không rõ ràng và mơ hồ. Ví dụ, liệu “lòng từ bi” (compassion) có tương tự như “sự đồng cảm” (empathy) hay “sự thương cảm” (sentimentality)? Nếu bạn có một luận điểm nói rằng “*lòng từ bi liên quan tích cực đến sự*

đồng cảm”, bạn không thể kiểm định luận điểm này trừ khi bạn phân biệt rõ ràng khái niệm “lòng từ bi” và khái niệm “sự đồng cảm”, sau đó đánh giá một cách chính xác bằng thực nghiệm hai phạm trù rất giống nhau này. Nếu những người sùng đạo tin rằng một số thành viên của xã hội, chẳng hạn như những người vô thần, người đồng tính và các bác sĩ làm nhiệm vụ nạo phá thai, sẽ bị hoả thiêu trong địa ngục vì tội lỗi của họ, những con chiên này ngày đêm cố gắng thay đổi hành vi “tội lỗi” để tránh phải bị đày ải ở địa ngục. Vậy thì hành động của họ là đáng thương hay đáng trách? Định nghĩa về các phạm trù như vậy không có bất kỳ tiêu chuẩn khách quan nào mà dựa trên quan niệm chủ quan của chúng ta về phạm trù đó.

Trong khi định nghĩa các phạm trù như “định kiến” hay “lòng từ bi”, cần nhớ rằng đôi khi, các phạm trù này là không có thật hoặc có khả năng tồn tại riêng rẽ, nhưng chúng được sáng tạo trong nhận thức của chúng ta. Ví dụ, có thể có một số bộ lạc trên thế giới họ không có định kiến với bất kỳ vấn đề gì; do đó họ không thể tưởng tượng được khái niệm này là gì. Nhưng trong nghiên cứu, cần phải xem các khái niệm này là có thật. Quá trình coi những phạm trù tinh thần có thật được gọi là vật thể hóa (reification). Đó là cơ sở để định nghĩa các phạm trù và xác định các biến đo lường để đánh giá chúng.

Một nhiệm vụ quan trọng khi khái niệm hóa các phạm trù đó là phải xác định chúng là đơn hay phức. Phạm trù đơn chỉ có một nghĩa và có thể đo lường, đánh giá chỉ bằng một kiểm định, ví dụ cân nặng của người, tốc độ gió. Trong khi đó, phạm trù phức chứa đựng hai hay nhiều khía cạnh. Ví dụ, khi khái niệm hóa phạm trù “năng lực học tập cá nhân”, phạm trù này bao gồm khả năng toán học và khả năng ngôn ngữ; vậy nên mỗi khía cạnh cần phải được đánh giá một cách riêng biệt, sau đó kết quả sẽ được tổng hợp để đưa ra kết luận cuối cùng về năng lực học tập của một người.

Thao tác hóa

Sau khi một phạm trù được định nghĩa, câu hỏi đặt ra là làm thế nào để đánh giá một cách chính xác nó. Thao tác hóa (operationalization) nói đến quá trình phát triển các chỉ số (indicator) để đánh giá, đo lường các phạm trù này. Ví dụ, một phạm trù trừu tượng không thể quan sát được như phạm trù “thực trạng kinh tế - xã hội” được định nghĩa là mức thu nhập của hộ gia đình, thì quá trình thao tác hóa được thực hiện bằng cách sử dụng chỉ số “thu nhập hộ gia đình” và hỏi người tham gia: thu nhập hàng năm của gia đình bạn là những khoản gì? Bởi vì tính chủ quan và thiếu rõ ràng của các phạm trù trong khoa học xã hội (trừ một số phạm trù nhân khẩu học như tuổi, giới tính, bằng cấp, thu nhập), thường

đòi hỏi phải dùng nhiều chỉ số khác nhau để đánh giá một phạm trù. Quá trình này cho phép kiểm tra sự chặt chẽ của các chỉ số và coi đó như là một tiêu chí để đánh giá tính xác thực của chúng (độ tin tưởng).

Nếu như các phạm trù được định nghĩa về mặt lý thuyết thì các chỉ số lại được áp dụng trên thực tế. Sự kết hợp các chỉ số ở cấp độ thực nghiệm được gọi là một biến số. Như đã đề cập trong chương trước, biến số có thể độc lập, phụ thuộc, điều hòa hay trung gian phụ thuộc vào cách chúng ta sử dụng chúng trong một nghiên cứu cụ thể. Bên cạnh đó, mỗi chỉ số có thể có một hoặc nhiều *lượng biến* (attribute) và mỗi lượng biến đại diện cho một *trị số* (value). Ví dụ, biến số “giới tính” có hai lượng biến là nam hoặc nữ. Tương tự, mức độ hài lòng của khách hàng có thể được phân chia thành năm lượng biến “rất không hài lòng”, “không hài lòng”, “không rõ”, “hài lòng” và “rất hài lòng”. Trị số của các lượng biến có thể là định lượng (bằng số) hay định tính (không bằng con số).

Dữ liệu định lượng có thể được phân tích bằng các kỹ thuật phân tích định lượng như hồi quy hay mô hình phương trình cấu trúc, trong khi đó dữ liệu định tính đòi hỏi các kỹ thuật định tính ví dụ như mã hóa (*coding*). Cần chú ý, nhiều biến trong khoa học xã hội là định tính, dù cho chúng được trình bày dưới hình thức định lượng. Ví dụ, chúng ta có thể tạo ra một chỉ số hài lòng của khách hàng với năm mức độ nêu trên và đánh số lần lượt từ 1 - 5 cho các mức độ đó. Vì vậy, có thể sử dụng các công cụ thống kê hiện đại để phân tích dữ liệu định lượng. Tuy nhiên, các con số này chỉ là những mào giúp người trả lời đánh giá mức độ hài lòng của họ và biến số “sự hài lòng” vẫn là định tính cho dù chúng được trình bày theo cách định lượng.

Có hai loại chỉ số là chỉ số phản ánh và chỉ số thành phần. Chỉ số *phản ánh* (reflective indicator) được dùng để diễn đạt bản chất của phạm trù. Ví dụ, nếu “mộ đạo” được định nghĩa là một phạm trù để đánh giá mức độ sùng đạo của một người thì việc tham dự các nghi lễ tôn giáo có thể là một chỉ số phản ánh của phạm trù mộ đạo. Chỉ số *thành phần* (formative indicator) dùng để hình thành hay góp phần hình thành nên phạm trù. Chỉ số này có thể đại diện cho các khía cạnh khác nhau của phạm trù. Ví dụ, nếu “mộ đạo” được định nghĩa là sự cấu thành bởi các góc độ: niềm tin, sùng bái và lễ nghĩa tôn giáo thì các chỉ số được chọn để đo lường các góc độ khác nhau này được coi là chỉ số thành phần. Phạm trù đơn được đánh giá bằng các chỉ số phản ánh, trong khi đó các phạm trù phức được đo lường bởi sự kết hợp các chỉ số thành phần và mỗi góc độ có thể được đo bằng cách sử dụng một hay nhiều chỉ số phản ánh.

Mức độ đánh giá

Công việc đầu tiên cần thực hiện khi thao tác hóa một phạm trù đó là quyết định các mức độ đánh giá cụ thể (level of measurement). Mức độ đánh giá (hay còn gọi là thang tỉ lệ - rating scale) đề cập đến các giá trị của một chỉ số. Ví dụ, nam và nữ (hoặc M và F hoặc 1 và 2) là hai giá trị của chỉ số “giới tính”. Trong bài viết chuyên đề với tựa đề "Lý thuyết về thang đánh giá" được xuất bản trong tạp chí “Science” vào năm 1946, nhà tâm lý học Stanley Smith Stevens (1946) đã xác định bốn loại thang đánh giá cơ bản dùng cho đánh giá khoa học: thang định danh, thang thứ bậc, thang khoảng và thang tỷ lệ. Đặc điểm thống kê của các thang này được thể hiện trong Bảng 6.1.

Thang	Xu hướng trung tâm	Kỹ thuật thống kê	Sự biến đổi
Định danh	Yếu vị (Mode)	Kiểm định Ki bình phương (Chi-square)	Lần lượt (bình đẳng)
Thứ bậc	Trung vị	Phân vị, thống kê phi tham số	Tăng dần (tuần tự)
Khoảng	Trung bình cộng, khoảng biến thiên, độ lệch chuẩn	Tương quan, hồi quy, phân tích sự khác biệt	Tuyến tính dương (affine)
Tỷ lệ	Trung bình hình học, trung bình điều hoà	Hệ số biến thiên	Đồng dạng dương (phép nhân, phép lôga)

Chú ý: Các thang cao có thể sử dụng tất cả kỹ thuật thống kê của thang thấp

Bảng 6.1. Thuộc tính thống kê của các thang tỉ lệ

Thang định danh (nominal scale), cũng được gọi là thang xác thực, dùng để đo lường, đánh giá sự xác thực của các nguồn dữ liệu. Những thang này được sử dụng cho những biến hay những chỉ số có thành phần (hay giá trị) không loại bỏ lẫn nhau, ví dụ như giới tính (hai giá trị: nam hoặc nữ), ngành công nghiệp (sản xuất vật liệu, tài chính, nông nghiệp,...) hay tôn giáo (Thiên Chúa, Đạo Hồi, Do Thái,...). Kể cả khi chúng ta đặt số cho mỗi giá trị (ví dụ 1 cho nam và 2 cho nữ) thì những con số này cũng không thật sự mang đến ý nghĩa gì cả (ví dụ 1 không có nghĩa là nhỏ hơn hay là một nửa của 2) và chúng cũng dễ dàng được đặt tên, ví như nam là M và nữ là F. Các thang định danh chỉ đơn thuần cung cấp tên hay nhãn mác của các giá trị khác nhau. Đo lường xu hướng trung tâm của một thang định danh là mode (yếu vị), cả trung vị hay trung bình đều không được định nghĩa. Kỹ thuật chi-square (xác suất) và phân bố tần đều có thể được sử dụng và chỉ cho phép thực hiện sự biến đổi lần lượt (bình đẳng), ví dụ 1 là nam, 2 là nữ.

Thang đo thứ bậc (ordinal scale) đo lường những dữ liệu đã được sắp xếp theo thứ tự, chẳng hạn như thứ bậc của sinh viên trong lớp học là thứ nhất, thứ hai, thứ ba và thứ tư tùy thuộc vào điểm trung bình học tập của họ. Tuy nhiên, giá trị thực tế hoặc giá trị tương đối của các thành phần hay sự khác biệt trong giá trị của các thành phần không thể được đánh giá. Ví dụ, thứ hạng của các học sinh trong lớp không nói lên điểm trung bình cụ thể của họ. Một ví dụ điển hình trong khoa học tự nhiên là thang về độ cứng khoáng sản của Moh, trong đó mức độ cứng của các khoáng chất khác nhau được xác định bởi khả năng làm xước các khoáng sản khác. Ví dụ, kim cương có thể làm xước tất cả các khoáng sản khác xuất hiện trên trái đất, do đó kim cương là khoáng chất “cứng nhất”. Tuy nhiên, loại thang này không chỉ ra độ cứng cụ thể của các khoáng chất là bao nhiêu và cũng không cung cấp những đánh giá tương đối về độ cứng của chúng.

Thang thứ bậc cũng có thể sử dụng các nhãn mức độ như “xấu”, “trung bình” và “tốt”, hay “rất không hài lòng”, “không hài lòng”, “không biết”, “hài lòng” và “rất hài lòng”. Trong trường hợp thứ hai, chúng ta có thể nói rằng những người trả lời là “hài lòng” sẽ ít thỏa mãn hơn những người chọn “rất hài lòng” nhưng chúng ta không định lượng chính xác được mức độ hài lòng của họ. Xu hướng trung tâm của một thang thứ bậc có thể là yếu vị hoặc trung vị, trong khi đó trung bình không thể phát huy tác dụng. Do đó, phân tích thống kê có thể bao gồm phân tích phân vị và phi tham số, nhưng các kỹ thuật phức tạp hơn như hồi quy, tương quan và phân tích phương sai lại không phù hợp. Chuyển đổi tăng dần đều (mà vẫn giữ được xếp hạng) được phép áp dụng.

Thang đo thứ bậc không chỉ đánh giá theo trật tự thứ hạng, mà còn chỉ ra các khoảng bằng nhau giữa các chỉ số liền kề. Ví dụ, thang nhiệt độ (F hoặc C) có sự khác biệt trong khoảng 30 - 40 độ F giống như sự khác biệt trong khoảng 80 và 90 độ F. Tương tự như vậy, thang thu nhập hàng năm của một người có thể sử dụng các khoảng chỉ số: \$ 0 đến 10.000, \$ 10.000 đến 20.000, \$ 20.000 đến 30.000,... Đây là thang khoảng bởi vì điểm giữa của mỗi khoảng (ví dụ, \$ 5.000, \$ 15.000, \$ 25.000,...) cách đều nhau. Thang khoảng cho phép chúng ta đánh giá mức độ chênh lệch giữa các chỉ số. Đây là điểm mạnh mà khoảng định danh và khoảng thứ bậc không có được.

Xu hướng trung tâm chấp nhận được bao gồm trung bình, trung vị và yếu vị. Các phân tích thống kê có thể sử dụng bao gồm tất cả kỹ thuật trong thang định danh và thứ bậc cùng với các kỹ thuật tương quan, hồi quy, phân tích phương sai,... Chuyển đổi thang chấp nhận được là tuyến tính tích cực. Cần lưu ý rằng các thang về sự hài lòng được thảo luận trước đó là không hoàn toàn là một thang khoảng, bởi vì chúng ta không thể khẳng

định liệu khoảng khác biệt giữa "rất hài lòng" và "hài lòng" có tương tự như những khoảng khác biệt giữa "không biết" và "hài lòng" hay giữa "không hài lòng" và "rất không hài lòng" hay không. Tuy nhiên, các nhà nghiên cứu khoa học xã hội thường chấp nhận (dù là không chính xác) những khác biệt này là tương tự. Vì vậy, họ có thể sử dụng các kỹ thuật thống kê để phân tích dữ liệu.

Thang tỉ lệ (ratio scale) là thang chứa đựng tất cả đặc điểm của các thang định danh, thứ bậc và tỉ lệ, ngoài ra còn có gốc "0 có ý nghĩa" (true zero) (nói đến việc thiếu vắng hoặc không sẵn có của các phạm trù cơ bản). Hầu hết các hệ thống đo lường trong khoa học tự nhiên và kỹ thuật, chẳng hạn như khối lượng, mặt phẳng nghiêng, điện lượng,... và một số biến trong khoa học xã hội như tuổi tác, nhiệm kỳ trong tổ chức và quy mô doanh nghiệp (số lượng nhân viên hay tổng doanh thu) đều có sử dụng thang tỷ lệ. Ví dụ, một công ty có quy mô là "không" có nghĩa là nó không có nhân viên hoặc các khoản thu. Tất cả kỹ thuật của xu hướng trung tâm, bao gồm cả phương tiện hình học và điều hòa, tất cả các phương pháp thống kê và các chuyển đổi phức tạp như đồng vị dương được phép sử dụng trong thang tỷ lệ.

Dựa trên bốn kiểu loại thang được trình bày trên, chúng ta có thể hình thành những thang đánh giá trong khoa học xã hội. Trong đó, phổ biến nhất là thang nhị phân, thang Likert, thang đối nghĩa và thang Guttman.

Thang nhị phân (binary scale) là những thang định danh chứa đựng những mục chứa hai trị số ví dụ như có/không hay đúng/sai. Phạm trù "tích cực chính trị" là một thang nhị phân tiêu biểu, nó chứa 6 điểm nhị phân (Bảng 6.2). Mỗi điểm trong thang này là một mục nhị phân và tổng số những câu trả lời "đã từng" có thể được sử dụng như một thông số tổng thể để đánh giá sự tích cực chính trị của người đó. Thang nhị phân cũng có thể được sử dụng cho các biến khác, chẳng hạn giới tính (nam hay nữ), tình trạng việc làm (toàn thời gian hay bán thời gian),... Nếu mục "tình trạng việc làm" được sửa đổi để tạo ra nhiều hơn hai giá trị (ví dụ, thất nghiệp, toàn thời gian, bán thời gian và đã nghỉ hưu), nó không còn là nhị phân nhưng vẫn là một mục định danh.

Đã bao giờ bạn viết thư gửi cơ quan công quyền	Đã từng	Chưa từng
Đã bao giờ bạn ký bản kiến nghị liên quan đến chính trị	Đã từng	Chưa từng
Đã bao giờ bạn quyên góp tiền cho một hoạt động vì mục đích chính trị	Đã từng	Chưa từng
Đã bao giờ bạn quyên góp tiền cho một ứng viên tranh cử vào cơ quan công quyền	Đã từng	Chưa từng
Đã bao giờ bạn viết thư phản ánh vấn đề chính trị gửi ban biên tập một tờ báo hoặc tạp chí	Đã từng	Chưa từng
Đã bao giờ bạn thuyết phục ai đó thay đổi ý định bầu cử của họ	Đã từng	Chưa từng

Bảng 6.2. Thang nhị phân 6 điểm để đánh giá sự tích cực chính trị

Thang Likert (Likert scale) được thiết kế bởi Rensis Likert, là một thang đánh giá rất thông dụng trong khoa học xã hội cho dữ liệu định danh. Thang này chứa những mục là những phát biểu đơn giản, dễ hiểu giúp người trả lời bày tỏ mức độ đồng ý hay không đồng ý về một thang 5 đến 7 điểm bắt đầu từ rất đồng ý đến rất không đồng ý. Thang Likert là thang tổng bởi vì điểm của thang là một tổng các giá trị chỉ số của mỗi điểm khi được chọn bởi người trả lời.

	Hoàn toàn không đồng ý	Không đồng ý	Không ý kiến	Đồng ý	Hoàn toàn đồng ý
Tôi cảm thấy hài lòng với công việc của mình	1	2	3	4	5
Tôi có quan hệ tốt với đồng nghiệp	1	2	3	4	5
Tôi hạnh diện về mối quan hệ của tôi với người phụ trách công việc	1	2	3	4	5
Tôi có thể nói rằng mọi người vui lòng về sự có mặt của tôi tại cơ quan	1	2	3	4	5
Tôi có thể nói rằng đồng nghiệp tôn trọng tôi	1	2	3	4	5
Tôi cảm thấy tôi có những đóng góp thiết thực cho công việc	1	2	3	4	5

Bảng 6.3. Thang Likert 6 điểm để đánh giá mức độ hài lòng trong công việc

Thang Likert mang đến được độ chi tiết cao hơn (câu trả lời linh động hơn) so với thang nhị phân khi nó bao gồm cả câu trả lời trung lập. Chúng ta cũng có thể được sử dụng ba hoặc chín trị số (thường được gọi là các “mấu” - anchor), nhưng nên sử dụng số lượng trị số là lẻ để có thể tạo ra trị số trung lập. Một số nghiên cứu sử dụng “*lựa chọn cưỡng bức*” để buộc người trả lời đồng ý hay không đồng ý với các phát biểu bằng cách loại bỏ mấu trung lập và sử dụng lượng trị số chẵn. Tuy nhiên điều này không phải là một chiến lược tốt vì một số người thực sự có quan điểm trung lập và kỹ thuật này không cho họ cơ

hội để phát biểu quan điểm thật sự. Một đặc tính quan trọng của thang Likert là mặc dù các phát biểu là khác nhau ở các trị số khác nhau nhưng các mẫu “hoàn toàn đồng ý” và “hoàn toàn không đồng ý” là luôn luôn giống nhau. Thang Likert là thang thứ bậc bởi vì các mẫu không luôn luôn là các khoảng cách bằng nhau, tuy vậy, đôi khi chúng ta coi nó như thang khoảng.

Thang đối nghĩa (Semantic Differential Scale) là thang chứa nhiều chỉ số đối xứng về nghĩa, được sử dụng để tìm hiểu quan điểm của người được hỏi về một phát biểu nào đó bằng cách sử dụng các tính từ được đặt ở những vị trí đối xứng nhau. Ví dụ, phạm trù “thái độ về bảo hiểm y tế” có thể được đánh giá bằng cách sử dụng bốn cặp chỉ số như được minh họa trong Bảng 6.4. Giống như trong thang Likert, số điểm tổng của thang được tổng hợp từ điểm của các chỉ số thành phần. Tuy nhiên, cần chú ý rằng trong thang Likert, phát biểu có thể thay đổi nhưng sử dụng các trị số giống nhau để đánh giá mỗi phát biểu. Trong khi đó, ở thang đối nghĩa, một phát biểu có thể được đánh giá bởi các trị số khác nhau. Thang đối nghĩa được đánh giá là một công cụ rất thích hợp và hiệu quả để đo lường, đánh giá thái độ hay cảm xúc của con người về các đối tượng, các sự kiện hoặc các hành vi.

Bạn cho biết ý kiến của bạn về bảo hiểm y tế?						
	Rất tán thành	Tán thành	Không ý kiến	Tán thành	Rất tán thành	
Tốt	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Kém
Hữu ích	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Vô ích
Chu đáo	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Thiếu chu đáo
Thú vị	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Nhàm chán

Bảng 6.4. Thang đối nghĩa để đo thái độ về bảo hiểm y tế

Thang Guttman (Guttman scale) là thang đối xứng được thiết kế bởi Louis Guttman. Nó sử dụng hàng loạt những chỉ số được sắp xếp theo thứ tự tăng dần về mức tăng dần về cường độ của phạm trù cần đánh giá. Ví dụ, phạm trù “thái độ đối với người nhập cư” có thể được đánh giá bằng cách sắp đặt năm chỉ số như được minh họa trong Bảng 6.5. Mỗi chỉ số trong thang Guttman thể hiện một trọng số khác nhau (chưa được trình bày ở trên) thay đổi theo cường độ của chỉ số đó và sự kết hợp trọng số của mỗi câu trả lời được sử dụng như biện pháp tổng hợp của một quan sát.

Bạn đánh giá quan điểm của mình về người nhập cư trong các phát biểu dưới đây?		
Bạn cảm thấy phiền nếu người nhập cư trở thành công dân nước bạn	Có	Không
Bạn cảm thấy phiền nếu người nhập cư sống trong khu dân cư của bạn	Có	Không
Bạn cảm thấy phiền nếu sống cạnh nhà người nhập cư	Có	Không
Bạn cảm thấy phiền nếu có một người nhập cư là bạn thân	Có	Không
Bạn cảm thấy phiền nếu có người trong gia đình bạn cưới một người nhập cư	Có	Không

Bảng 6.5. Thang Guttman năm điểm để đánh giá thái độ với người nhập cư

Lập thang đánh giá (Scaling)

Trong phần trước, chúng ta đã tìm hiểu về các phương pháp khác nhau để đo lường, đánh giá chính xác những quan điểm, thái độ của người tham gia nghiên cứu thông qua việc sử dụng các phiếu khảo sát có chứa hàng loạt các chỉ số thường được trình bày dưới dạng các phát biểu hoặc các câu hỏi kèm theo các câu trả lời tự chọn. Nhưng làm thế nào để tạo ra các chỉ số đó? Quá trình tạo ra các chỉ số được gọi là lập thang đo lường hay lập thang đánh giá (scaling). Cụ thể hơn, lập thang đánh giá là một bước trong quá trình thực hiện việc đo lường, đánh giá các khái niệm, phạm trù, có nhiệm vụ xây dựng các chỉ số thông qua quá trình chuyển đổi các nội dung định tính trong khái niệm trừu tượng thành các chỉ số đo lường định lượng cụ thể. Stevens (1946) cho rằng: “Lập thang đánh giá là quá trình chuyển đổi các khái niệm nghiên cứu thành các con số theo một quy tắc nhất định”. Trên thực tế, quá trình chuyển đổi này luôn là một trong những nhiệm vụ khó khăn nhất khi tiến hành các nghiên cứu thực nghiệm trong các ngành khoa học xã hội.

Thành quả của quá trình lập thang (hay còn gọi là *lên thang*) là tạo ra được các thang đánh giá có chứa các chỉ số vừa có khả năng phản ánh toàn diện, khách quan nội dung bản chất của khái niệm, phạm trù trong nghiên cứu, vừa có thể kiểm nghiệm được trong thực tế. Như vậy, thang đánh giá (rating scales) là sản phẩm của lập thang đánh giá (scaling). Thang đánh giá có nhiều loại khác nhau, được sử dụng để nắm bắt ý kiến của người trả lời thông qua các chỉ số cụ thể. Ví dụ như thang định danh sẽ xác định ý kiến đồng ý hay không đồng ý; thang khoảng sẽ xác định được mức độ đồng ý bằng việc sử dụng hàng loạt giá trị từ "rất không đồng ý" đến "rất đồng ý". Việc hình thành các chỉ số với các giá trị kèm theo này chính là quá trình lập thang đánh giá.

Như đã đề cập ở phần trước, các khái niệm và phạm trù cần đo lường có thể là đơn (ví dụ: cân nặng, tốc độ gió, quy mô một tổ chức) hoặc phức (ví dụ: năng lực học thuật, trí thông minh), vì vậy, các thang đánh giá cũng có thể đơn hoặc phức. Thang đơn đo lường các khái niệm bằng một thang duy nhất, với các chỉ số kèm theo có giá trị từ cao xuống thấp. Chú ý là các thang đơn này có thể sử dụng nhiều chỉ số khác nhau, nhưng tất cả chúng đều hướng tới việc đánh giá, đo lường cùng một khái niệm hoặc một phạm trù. Trong khoa học xã hội, thang đơn phù hợp với các phạm trù đơn phản ánh một khía cạnh với các mức độ từ thấp đến cao, ví dụ như phạm trù sự tự trọng.

Trong khi đó, thang phức sử dụng nhiều chỉ số khác nhau để đánh giá một cách riêng biệt từng khía cạnh của phạm trù phức, sau đó kết hợp điểm số từ các khía cạnh đó để đưa ra đánh giá tổng quát về phạm trù. Ví dụ phạm trù năng lực học thuật có thể được đánh giá bằng cách đo lường một cách riêng biệt hai khía cạnh là năng lực nghiên cứu và năng lực diễn đạt, sau đó kết hợp điểm số của hai khía cạnh này sẽ được kết quả chung cho phạm trù năng lực học thuật. Bởi vì phần lớn thang đánh giá được sử dụng trong khoa học xã hội là các thang đơn, do vậy ở phần tiếp theo, chúng ta sẽ tìm hiểu ba cách thức khác nhau để hình thành nên các thang đơn.

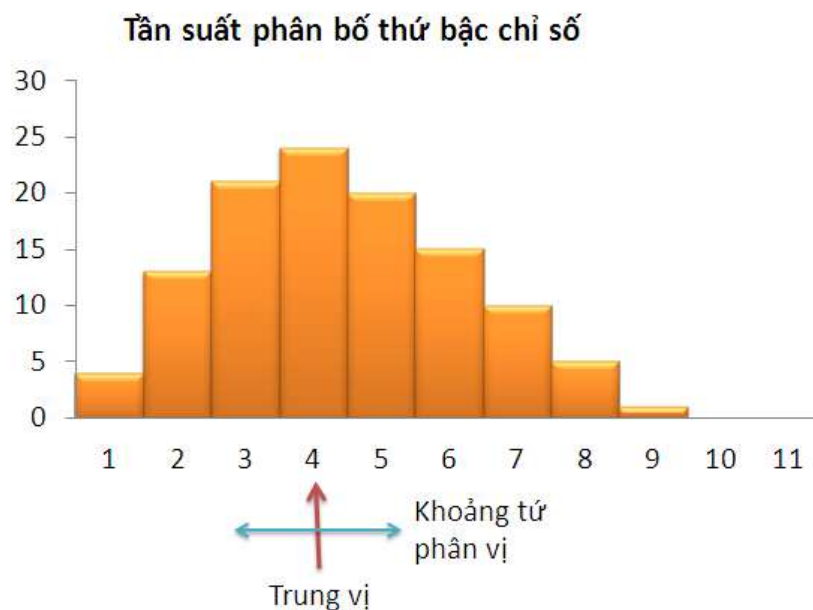
Các phương pháp lập thang đơn được hình thành trong nửa đầu thế kỷ XX và mang tên những người tạo ra. Ba phương pháp lập thang đơn phổ biến nhất là: (1) Phương pháp lập thang đồng đều của Thurstone, (2) Phương pháp lập thang tổng hợp của Likert, (3) Phương pháp lập thang lũy kế của Guttman. Ba cách tiếp cận này có sự tương đồng trên nhiều phương diện, nhưng cũng có sự khác nhau cơ bản ở tỷ lệ các chỉ số theo các phương pháp xem xét và thống kê thường dùng để lựa chọn. Mỗi phương pháp này sẽ được thảo luận ở các phần tiếp theo.

Phương pháp lập thang đồng đều của Thurstone (Thurstone's equal-appearing scaling method). Louis Thurstone, một trong những nhà lý luận đầu tiên và nổi tiếng nhất về lập thang đánh giá. Năm 1925, ông công bố phương pháp về lập thang đồng đều. Phương pháp này bắt đầu với việc đưa ra một định nghĩa rõ ràng về phạm trù nghiên cứu. Dựa trên định nghĩa này, các chỉ số tiềm năng phản ánh bản chất phạm trù quan tâm sẽ được liệt kê. Những chỉ số này được gợi ý bởi các chuyên gia có chuyên môn về phạm trù cần đo lường. Một loạt chỉ số tiềm năng (lý tưởng là khoảng từ 80 - 100 chỉ số) được trình bày theo một hình thức thống nhất, ví dụ như sử dụng hình thức là các phát biểu để người tham gia nghiên cứu tích vào là đồng ý hay không đồng ý.

Sau đó một ban thẩm định được tuyển chọn để quyết định sẽ chọn các chỉ số cụ thể nào trong số các chỉ số tiềm năng để sử dụng trong phiếu khảo sát. Ban thẩm định có thể bao gồm những chuyên gia có chuyên môn trong việc xây dựng, sử dụng các công cụ đo lường hoặc có thể là một mẫu ngẫu nhiên những người sẽ tham gia trả lời phiếu khảo sát. Mỗi thành viên ban thẩm định dựa trên quan điểm cá nhân của họ sẽ độc lập đánh giá mức độ ủng hộ, đồng ý với nội dung phạm trù của mỗi chỉ số (dưới dạng một phát biểu) theo thang điểm từ 1 tới 11 (1 có nghĩa là họ nghĩ rằng phát biểu này rất không sát hợp không đồng ý với nội dung của phạm trù và 11 nghĩa là nó rất đồng ý).

Bước tiếp theo là tính toán giá trị trung bình và khoảng giá trị tứ phân vị (còn gọi là giá trị một phần tư) của mỗi chỉ số. Khoảng giá trị tứ phân vị là hiệu số giữa giá trị tứ phân vị thứ ba (phần trăm thứ 75) và giá trị tứ phân vị thứ nhất (phần trăm thứ 25) trong tổng số 100% tổng giá trị của mỗi chỉ số. Hình 6.1 biểu thị giá trị trung bình và khoảng giá trị tứ phân vị của một chỉ số. Cuối cùng, dựa trên các khoảng bằng nhau trong phạm vi các giá trị trung bình, một vài chỉ số nhất định sẽ được lựa chọn. Điều này có thể được tiến hành bằng cách chọn các chỉ số có cùng một giá trị trung bình, sau đó có thể chọn thêm các chỉ số mà khoảng tứ phân vị của chúng có cách biệt nhỏ nhất so với giá trị trung bình của các chỉ số được chọn trước. Cũng nên lưu ý, thay vì chỉ hoàn toàn dựa vào các tính toán số học, cũng cần phải xem xét đến ý nghĩa và sự rõ ràng của các chỉ số (các phát biểu). Giá trị trung bình của mỗi chỉ số sẽ làm cơ sở cho việc tổng hợp một thang chung để đánh giá phạm trù nghiên cứu. Lúc này, chúng ta có một thang đánh giá trong giống như một cái thước kẻ có 11 điểm với các chỉ số có giá trị trung bình khác nhau được đặt ở vị trí khác nhau của thước. Bởi các chỉ số xuất hiện một cách đồng đều chạy dọc 11 điểm của thang, nên kỹ thuật này được gọi là lập thang xuất hiện đồng đều (equal-appearing scaling method).

Ngoài ra, Thurstone cũng đưa ra hai phương pháp khác để xây dựng các thang đơn, đó là phương pháp khoảng liên tục (method of successive intervals) và phương pháp so sánh cặp (method of paired comparisons). Cả hai phương pháp này có nhiều điểm tương tự như phương pháp khoảng đồng đều, chỉ khác nhau ở cách chúng ta yêu cầu các thành viên ban thẩm định đưa ra đánh giá của họ về các chỉ số. Ví dụ, trong xây dựng thang so sánh cặp, mỗi thẩm định viên được yêu cầu đưa ra đánh giá của họ cho từng cặp chỉ số (chứ không phải là đánh giá mỗi một chỉ số theo thang 1 đến 11). Trên thực tế, đối với những nghiên cứu có nhiều chỉ số, phương pháp này sẽ rất mất thời gian và có tiến độ chậm hơn nhiều so với phương pháp khoảng đồng đều.



Hình 6.1. Biểu đồ các giá trị của chỉ số thang Thurstone

Phương pháp lập thang tổng hợp của Likert (Likert’s summative scaling method).

Phương pháp lập thang tổng hợp của Likert, một phương pháp lập thang đơn do Murphy và Likert (1938) phát triển, có thể là phương pháp phổ biến nhất trong số ba phương pháp lập thang được giới thiệu trong chương này. Cũng giống như phương pháp của Thurstone, phương pháp lập thang của Likert bắt đầu với một định nghĩa rõ ràng về phạm trù nghiên cứu. Sau đó sử dụng một nhóm các chuyên gia để tạo ra khoảng từ 80 đến 100 chỉ số tiềm năng. Những chỉ số này sẽ được các thẩm định viên đánh giá và được phân vào một trong số năm nhóm (hoặc 1 đến 7 nhóm) sau: (1) là hoàn toàn không đồng ý với nội dung phạm trù, (2) phần nào không đồng ý, (3) chưa quyết định, (4) phần nào đồng ý, (5) hoàn toàn đồng ý.

Sau khi phân loại, việc lựa chọn các chỉ số cuối cùng cho thang đánh giá có thể thông qua một trong số cách sau: (1) tính toán các tương quan lưỡng biến giữa mỗi chỉ số riêng lẻ với chỉ số tổng (là tổng số các chỉ số đưa ra cho người trả lời), sau đó loại bỏ các chỉ số có tương quan thấp (ví dụ nhỏ hơn 0.60), (2) tính giá trị đánh giá trung bình cho mỗi chỉ số, sử dụng phương pháp t-test để tính độ lệch giữa các giá trị trung bình đó và sau đó chọn những chỉ số có giá trị *t* (t-value) cao (đó là những chỉ số có độ lệch cao nhất giữa các giá trị tứ phân vị). Cuối cùng, nhà nghiên cứu xem xét, lựa chọn một số tương đối ít các chỉ số (ví dụ từ 10 đến 15) vừa có tương quan cao giữa chỉ số cá thể và chỉ số tổng và vừa có độ lệch cao (giá trị *t* cao). Phương pháp Likert giả định các chỉ số có trọng số giống nhau, vì vậy phản ứng của người tham gia nghiên cứu đối với mỗi chỉ số có thể được tổng hợp để

tạo ra một điểm số chung cho người tham gia đó. Vì lý do này, phương pháp này được gọi là lập thang tổng hợp. Cần chú ý, trước khi tổng hợp, những chỉ số có nghĩa đối ngược với nghĩa ban đầu của phạm trù nghiên cứu phải được đảo ngược lại (ví dụ 2 đảo thành 5, 2 thành 4,...).

Phương pháp lập thang lũy kế của Guttman (Likert's summative scaling method). Phương pháp lập thang lũy kế (hoặc lập thang lũy tiến hoặc tích lũy) được Guttman (1950) thiết kế dựa trên ý tưởng về khoảng cách xã hội của Emory Bogardus. Ý tưởng này giả định rằng sự nhiệt tình của mỗi cá nhân khi tham gia vào các mối quan hệ xã hội với người khác là khác nhau. Phương pháp Guttman cũng dựa trên ý tưởng về cường độ, sử dụng một danh sách chứa đựng các chỉ số được phân bố theo cường độ từ “ít kịch tính nhất” đến “kịch tính nhất”. Điều đó có nghĩa là nếu ai đó đồng ý với một chỉ số trong danh sách, họ sẽ đồng ý với tất cả các chỉ số có cường độ thấp hơn (ít kịch tính hơn) được liệt kê trước đó. Tuy nhiên trên thực tế, chúng ta rất hiếm khi tìm được một bộ chỉ số thể hiện một cách trọn vẹn mô hình lũy kế này. Để giải quyết vấn đề này, một phân tích cấu trúc tiềm ẩn (scalogram analysis) cũng sẽ được sử dụng để kiểm tra mức độ tương ứng của bộ chỉ số với độ lũy kế của thang.

Cũng giống như các phương pháp lập thang trước, phương pháp Guttman cũng bắt đầu với một định nghĩa rõ ràng về phạm trù nghiên cứu, sau đó sử dụng các chuyên gia để hình thành một số lượng tương đối lớn các chỉ số tiềm năng. Một nhóm thẩm định viên sẽ đánh giá mỗi chỉ số tiềm năng này, sau đó phân loại là “có” nếu họ thấy chỉ số thể hiện sự đồng ý, tán thành với nội dung của phạm trù và phân loại là “không” nếu họ cho rằng chỉ số này không đồng ý. Tiếp theo, lập một bảng trong đó chỉ rõ kết quả phân loại của các thẩm định viên cho mỗi chỉ số. Các hàng ngang của bảng này được sắp xếp theo thứ tự giảm dần về số lượng các phân loại “có” của các thẩm định viên. Điều đó có nghĩa là thẩm định viên có nhiều phân loại “có” sẽ ở phần trên của bảng và ngược lại những người có ít phân loại “có” sẽ ở phần dưới. Để sắp xếp các cột dọc, các chỉ số được phân loại “có” nhiều hơn từ các thẩm định viên sẽ được xếp bên trái và giảm dần sang phải theo số lượng “có” của từng chỉ số (xem ví dụ về bảng này trong Bảng 6.6). Như vậy lúc này khi quan sát từ trái sang phải (đọc theo các chỉ số), thang đáng giá gần như được xem là lũy kế. Tuy vậy, có thể có một vài ngoại lệ làm cho thang không hoàn toàn là lũy kế như được phản ánh trong Bảng 6.6. Để quyết định đâu là các chỉ số phản ánh chính xác nhất đặc điểm lũy kế này, một kỹ thuật phân tích dữ liệu được gọi là phân tích cấu trúc tiềm ẩn có thể được sử dụng (trong trường hợp chỉ có một số nhỏ các chỉ số, bước này cũng có thể

được tiến hành một cách trực quan). Kỹ thuật phân tích này cũng dự kiến điểm số cho mỗi chỉ số, làm cơ sở để tính toán điểm số chung cho tất cả chỉ số của một thẩm định viên.

Thẩm định viên	Chỉ số 12	Chỉ số 5	Chỉ số 3	Chỉ số 22	Chỉ số 8	Chỉ số 7	...
29	Y	Y	Y	Y	Y	Y	
7	Y	Y	Y	-	<u>Y</u>	-	
15	Y	Y	Y	Y	-	-	
3	Y	Y	Y	Y	-	-	
32	Y	Y	Y	-	-	-	
4	Y	Y	-	<u>Y</u>	-	-	
5	Y	Y	-	-	-	-	
23	Y	Y	-	-	-	-	
11	Y	-	-	<u>Y</u>	-	-	
<u>Y</u> Trường hợp ngoại lệ cản trở khả năng luỹ kế hoàn hảo							

Bảng 6.6. Bảng đánh giá thang Guttman sau khi sắp xếp

Chỉ số so sánh

Chỉ số so sánh (index) là điểm số tổng hợp dựa trên việc kết hợp những đánh giá đối với nhiều phạm trù (được gọi là các thành phần) bằng cách sử dụng một bộ quy tắc và công thức. Cũng là một đánh giá tổng hợp nhưng chỉ số so sánh khác các thang ở chỗ chỉ số đánh giá các khía cạnh khác nhau của cùng một phạm trù. Ví dụ điển hình của chỉ số so sánh là chỉ số giá tiêu dùng (CPI), được tính toán hàng tháng bởi Cục Thống kê Lao động của Bộ Lao động Mỹ. Chỉ số CPI là thước đo để chỉ ra người tiêu dùng đã chi trả bao nhiêu cho việc mua sắm các loại hàng hóa và dịch vụ.

Chỉ số này được chia thành tám nhóm lớn: Thực phẩm và đồ uống, nhà ở, quần áo, giao thông đi lại, sức khỏe y tế, giải trí, học tập - thông tin liên lạc, các hàng hóa và dịch vụ khác. Những nhóm này được chia nhỏ thành hơn 200 nhóm hàng nhỏ hơn. Mỗi tháng, các nhân viên chính phủ trên cả nước Mỹ được đề nghị cung cấp mức giá hiện tại của hơn 80.000 mặt hàng khác nhau. Sử dụng một chương trình tính toán phức tạp cùng một loạt công thức, quy tắc để xem xét địa bàn và khả năng mua sắm của từng hạng mục hàng hóa, giá cả của các hàng hóa đó được kết hợp bởi các nhà phân tích để hình thành một chỉ số tổng thể.

Một ví dụ khác về chỉ số là hiện trạng kinh tế xã hội (socio-economic status - SES), cũng được gọi là chỉ số kinh tế xã hội (socio-economic index - SEI). Chỉ số này kết hợp ba phạm trù: thu nhập, giáo dục và nghề nghiệp. Thu nhập được đo bằng đô la, giáo dục đo bằng số năm đến trường hoặc bằng cấp đã đạt được và nghề nghiệp được phân loại thành các nhóm ngành nghề hoặc bằng vị trí công việc. Những số liệu này được kết hợp để tạo ra một chỉ số SES tổng bằng cách tổng hợp điểm số “giáo dục nghề nghiệp” (occupational education) – phần trăm số người trong một công việc đã có ít nhất một năm học tập tại môi trường đại học và “thu nhập nghề nghiệp” (occupational income) – phần trăm số người trong công việc đó đã hưởng lương ít nhất một năm.

Quá trình tạo ra một chỉ số so sánh cũng tương tự như đối với một thang. Đầu tiên cần phải định nghĩa khái niệm cho các chỉ số và các thành phần cấu thành của nó. Mặc dù điều này dường như đơn giản, nhưng có thể có rất nhiều bất đồng giữa khi xem xét những thành phần (cấu trúc) nào nên được bao gồm hoặc loại trừ khỏi ra một chỉ số. Ví dụ như trong chỉ số SES, liệu rằng thu nhập có mối tương quan với giáo dục và nghề nghiệp hay không, nếu có thì chúng ta chỉ cần thống kê một thành phần hay phải thống kê cả ba. Để giải quyết vấn đề này tốt nhất chúng ta cần phải khảo cứu những tài liệu, lý thuyết có liên quan và tham vấn những chuyên gia trong lĩnh vực lao động, việc làm.

Thứ hai, chúng ta cần thao tác hóa và sau đó đo lường mỗi thành phần. Ví dụ, bạn sẽ phân loại nghề nghiệp như thế nào, nhất là khi một số nghề có thể thay đổi theo thời gian (ví dụ, trước khi có Internet sẽ không có công việc phát triển Web). Thứ ba, tạo ra một quy tắc hay công thức để tính toán các chỉ số so sánh. Tương tự như các bước trước đó, việc hình thành các nguyên tắc và công thức cũng chứa đựng nhiều yếu tố chủ quan. Cuối cùng, tiến hành củng cố các chỉ số bằng cách sử dụng các dữ liệu hiện có hoặc dữ liệu mới.

Mặc dù chỉ số so sánh và thang đo lường mang lại các giá trị định lượng biểu thị cho một phạm trù đang nghiên cứu, chúng có những điểm khác nhau cơ bản. *Trước hết*, chỉ số thường bao gồm các thành phần rất khác nhau và được đo bằng nhiều cách khác nhau. Trong khi đó thang thường chứa đựng các chỉ số tương tự và sử dụng cùng một thang đo lường (ví dụ như một thang Likert năm điểm). *Thứ hai*, chỉ số so sánh thường kết hợp các chỉ số đo lường khách quan như giá cả, thu nhập; trong khi thang thường được thiết kế để đánh giá các phạm trù chủ quan hoặc có tính chất phán xét như thái độ, thành kiến hoặc lòng tự trọng. Một số người cho rằng sự tinh tế và phức tạp khi thiết kế các thang đo sẽ làm cho thang trở nên khác với chỉ số, trong khi đó, những người khác cho rằng phương thức

thiết kế chỉ số so sánh cũng tinh tế và phức tạp không kém. Tuy nhiên, rõ ràng cả chỉ số so sánh và thang đều là những công cụ thiết yếu trong nghiên cứu khoa học xã hội.

Mô hình hóa

Thang và chỉ số so sánh giúp tạo ra những đánh giá thứ tự về các phạm trù một chiều. Tuy nhiên, đôi khi các nhà nghiên cứu muốn tổng hợp kết quả đo lường của hai hay nhiều phạm trù nhằm tạo ra một mô hình (typology) phản ánh cơ chế hoạt động của một hệ đối tượng. Không giống như thang hay chỉ số so sánh, mô hình hóa là phức tạp, chỉ bao gồm các biến danh nghĩa. Ví dụ, một người có thể tạo ra một mô hình chính trị của các tờ báo dựa trên định hướng của chúng đối với chính sách đối nội và đối ngoại; được trình bày trong mục xã luận của những tờ báo này, như thể hiện trong Hình 6.2. Mô hình này có thể được sử dụng để phân loại báo chí thành một trong bốn “thể loại lý tưởng” (A đến D), xác định sự phân phối của báo chí tới các phân loại lý tưởng này.



Hình 6.2. Một cách phân loại khuynh hướng của báo chí

Tóm tắt

Nói tóm lại, việc xây dựng các thang, khoảng hay mô hình hóa trong nghiên cứu khoa học xã hội là một tiến trình phức tạp, bao gồm nhiều nhiệm vụ khác nhau. Một vài nhiệm vụ đó là:

- Bạn nên sử dụng thang, chỉ số hay mô hình hóa?
- Bạn lên kế hoạch để phân tích dữ liệu như thế nào?
- Phương thức đánh giá bạn mong muốn là gì (định danh, thứ tự, khoảng hay tỉ lệ)?
- Bạn nên sử dụng bao nhiêu chỉ số (ví dụ 1 đến 10, 1 đến 7, -3 tới +3)?

- Bạn nên sử dụng số lượng chỉ số chẵn hay lẻ (bạn có muốn có một trị số trung lập không?)

- Bạn đặt tên cho các chỉ số như thế nào (nhất là trong thang đối nghĩa)?

- Cuối cùng, lộ trình nào bạn sẽ sử dụng để tạo ra những thang và chỉ số so sánh đã chọn?

Chương này đã tìm hiểu tiến trình và sản phẩm của việc hình thành phương thức đánh giá các phạm trù. Độ tin cậy và giá trị khoa học của các phương thức đánh giá này sẽ được xem xét thấu đáo trong chương tiếp theo.

Chương 7

ĐỘ TIN CẬY VÀ GIÁ TRỊ CỦA THANG ĐÁNH GIÁ

Trong chương trước, chúng ta đã nhận diện những khó khăn gặp phải khi thiết lập các phương pháp đánh giá phạm trù trong nghiên cứu khoa học xã hội. Ví dụ, làm thế nào để có thể khẳng định chúng ta đang đánh giá phạm trù “lòng từ bi” mà không phải là phạm trù “sự đồng cảm” khi mà cả hai phạm trù này có những giao thoa về ngữ nghĩa. Nó còn phức tạp hơn với các phạm trù phức và chỉ tồn tại trong trí tưởng tượng (không tồn tại trên thực tế). Do đó, nếu chỉ đo lường, đánh giá các phạm trù trong khoa học xã hội bằng cách sử dụng các thang đo lường quen thuộc là không đầy đủ. Chúng ta cũng phải kiểm định những thang này để đảm bảo rằng: (1) chúng thực sự đánh giá các phạm trù mà chúng ta mong muốn (thang có “giá trị khoa học”), (2) chúng đánh giá các phạm trù mong muốn một cách liên tục và nhất quán (thang có “độ tin cậy”). Cùng với nhau, độ tin cậy và giá trị khoa học được xem là “linh hồn” của các phương thức đánh giá và là công cụ để xác định sự đúng đắn và đầy đủ của các phương pháp nghiên cứu khoa học.

Một phương pháp có thể đáng tin nhưng lại không có giá trị nếu như nó đo lường, đánh giá một đối tượng nào đó một cách liên tục và nhất quán nhưng lại đánh giá nhầm những đối tượng khác. Tương tự như vậy, một phương pháp có thể có giá trị nhưng không đáng tin nếu nó đánh giá chính xác những đối tượng cần nghiên cứu nhưng lại không theo một hướng liên tục. Sử dụng một minh họa về bia đỡ đạn như trong Hình 7.1 để có thể thấy một phương pháp đánh giá phạm trù phức vừa đáng tin cậy, vừa có giá trị nếu những viên đạn được bắn vào gần sát nhau quanh tâm của bia. Phương pháp có giá trị nhưng không đáng tin cậy nếu đạn được bắn xung quang tâm bia nhưng cách xa nhau. Cuối cùng, phương pháp đáng tin cậy nhưng không có giá trị nếu đạn bắn gần sát nhau nhưng lại không hướng vào tâm bia. Chính vì vậy, cả độ tin cậy lẫn giá trị khoa học đều cần thiết để đảm bảo phương pháp đánh giá các phạm trù cần nghiên cứu là đúng đắn.



Hình 7.1 So sánh độ tin cậy và giá trị khoa học

Độ tin cậy

Độ tin cậy (reliability) của một phương pháp là mức độ nhất quán hay đáng tin khi đánh giá một phạm trù. Nói cách khác, nếu chúng ta sử dụng một thang nào đó để đo lường, đánh giá cùng một phạm trù ở nhiều thời điểm khác nhau, thì liệu chúng ta có thu được kết quả tương tự trong mỗi lần đó và như vậy chúng ta có thể giả định rằng kết quả nghiên cứu không thay đổi? Ví dụ về một phương pháp đánh giá không đáng tin đó là hỏi người khác cân nặng của bạn. Khả năng những người khác nhau đưa ra những phán đoán khác nhau là khá phổ biến và những ước lượng đó là không nhất quán. Chính vì vậy, kỹ thuật ‘phán đoán’ là không đáng tin. Một phương pháp đánh giá có độ tin cậy cao hơn là khi dùng một cái máy cân để kiểm tra nhiều lần khác nhau và như vậy trọng lượng của bạn gần như không đổi trong mỗi lần cân, trừ khi cân nặng của bạn thật sự thay đổi giữa các lần đo lường đó.

Cần chú ý rằng độ tin cậy đề cập đến sự liên tục, nhất quán chứ không phải độ chính xác. Trong ví dụ trước về cân nặng, nếu như cái máy cân được sử dụng không đo một cách hoàn toàn chính xác (ví dụ nó giảm 10 pound so với khối lượng thực tế để làm cho người cân cảm thấy chút gì đó vui hơn). Do máy cân không đo lường trọng lượng thật của bạn, nên nó là một phương pháp không có giá trị khoa học.

Đâu là nguồn gốc của những quan sát thiếu tin cậy trong các phương pháp đánh giá của khoa học xã hội? Một trong số đó chính là sự chủ quan của người quan sát (nhà nghiên cứu). Nếu tâm lý, tinh thần của nhân viên trong một cơ quan hay tổ chức được đo bằng cách quan sát họ cười với nhau, hay họ nói chuyện đùa với nhau,... thì những người quan sát khác nhau có thể đưa ra những nhận định khác nhau về tinh thần của nhân viên khi họ quan sát trong một ngày vô cùng bận rộn của cơ quan (nhân viên không có thời gian để nói chuyện phiếm), trong khi số khác quan sát trong một ngày mà công việc nhẹ nhàng hơn, nhân viên vui vẻ và thích chuyện trò hơn. Thậm chí hai người quan sát cũng có thể đưa ra những đánh giá khác nhau về mức độ phấn chấn của nhân viên trong cùng một ngày nếu họ quan niệm thế nào là một câu chuyện phiếm. “Quan sát” là một kỹ thuật đánh giá định tính. Đôi khi độ tin cậy có thể được cải thiện, cũng có khi sử dụng các phương pháp định lượng, ví dụ xem xét việc đếm số lượng những lời phàn nàn và những lần xích mích trong một tháng là một phương pháp nghịch đảo để đánh giá tinh thần của nhân viên. Tất nhiên, số lần phàn nàn có thể hoặc không phải là một phương pháp có giá trị cao nhưng nó ít mang tính chủ quan và trở nên đáng tin hơn.

Nguyên nhân thứ hai của các quan sát thiếu tin cậy xuất phát từ những câu hỏi không rõ ràng và thiếu chính xác. Ví dụ, nếu bạn hỏi người khác thu nhập của họ là gì, những người được hỏi có thể hiểu khác nhau thế nào là thu nhập hàng tháng, hàng năm hay tính theo giờ. Điều này dẫn đến những kết quả rất khác nhau và thiếu độ tin tưởng. Nguyên nhân thứ ba đó là nội dung câu hỏi quá xa lạ, không liên quan đến người được hỏi, ví dụ như hỏi những sinh viên tốt nghiệp ở Mỹ về họ có hài lòng về mối quan hệ giữa Canada và Slovenia hay không hoặc giả đề nghị trưởng phòng tổ chức đánh giá hiệu quả của chiến lược công nghệ trong công ty - nội dung này nên hỏi trưởng phòng công nghệ.

Vậy thì làm cách nào bạn có thể tạo ra những phương pháp đáng tin? Nếu phương thức đánh giá của bạn cần đến sự cung cấp thông tin của bởi những người khác (đây là trường hợp hết sức phổ biến trong nghiên cứu khoa học xã hội) thì bạn có thể bắt đầu bằng cách thay thế các kỹ thuật thu thập dữ liệu phụ thuộc nhiều hơn vào sự chủ quan của nhà nghiên cứu (ví dụ phương pháp quan sát) bằng những kỹ thuật khách quan hơn (ví như điều tra bảng hỏi) và bằng cách hỏi những câu hỏi rõ ràng (ví dụ bạn đang muốn tìm hiểu về thu nhập hàng tháng) tới những người trả lời phù hợp (có hiểu biết liên quan). Những chiến thuật này có thể cải thiện độ tin cậy của phương pháp đánh giá bạn sử dụng, tuy vậy điều này không có nghĩa là phương pháp bạn chọn là hoàn toàn tin cậy. Những công cụ đánh giá phải được kiểm định về độ tin cậy của chúng. Như được thảo luận dưới đây, có nhiều cách khác nhau để xác định độ tin cậy.

Độ tin cậy nội bộ (Inter-rater reliability). Độ tin cậy nội bộ, còn được gọi là độ tin cậy giữa những người quan sát, được sử dụng để xác định sự nhất quán của hai hay nhiều người đánh giá (hay người quan sát) khi đánh giá cùng một phạm trù. Thông thường, nhiệm vụ này được thực hiện ở giai đoạn thử nghiệm của dự án nghiên cứu và có thể thực hiện theo hai cách, phụ thuộc và phương pháp đo lường. Đối với phương pháp pháp sử dụng thang phân loại thì phạm trù được phân loại thành các chỉ số khác nhau, sau đó những người đánh giá sẽ độc lập quan sát và liệt kê các kết quả quan sát vào các chỉ số tương ứng. Trung bình cộng số liệu tương ứng của những người quan sát sẽ phản ánh độ tin cậy nội bộ của thang đánh giá. Ví dụ có 2 người quan sát, đánh giá 100 kết quả quan sát đối với 3 chỉ số, nếu 75% kết quả quan sát của họ là giống nhau, thì độ tin cậy nội bộ là 0.75. Trong khi đó, đối với phương pháp đo lường sử dụng thang khoảng hoặc thang tỉ lệ (ví dụ có 2 người quan sát sử dụng thang khoảng từ 1 đến 7, mỗi lần quan sát hoạt động của lớp học cách nhau 5 phút), thì độ tin cậy nội bộ có thể được xác định bằng mức tương quan đơn giữa các đánh giá của những nhà quan sát.

Độ tin cậy liên tục (Test-retest reliability). Độ tin cậy liên tục dùng để xác định mức độ nhất quán của hai phương pháp đo lường cho cùng một phạm trù, với cùng một mẫu nhưng đo lường ở hai thời điểm khác nhau. Nếu như kết quả quan sát không có khác biệt lớn thì hai phương pháp này được xem là đáng tin. Độ tương quan trong các kết quả quan sát giữa hai lần đo lường sẽ là điểm số của độ tin cậy liên tục. Chú ý rằng khoảng thời gian giữa hai lần đo lường là cực kỳ quan trọng. Nhìn chung, khoảng thời gian càng dài thì khả năng các kết quả quan sát thay đổi càng cao và như vậy độ tin cậy liên tục sẽ càng thấp.

Độ tin cậy phân nửa (Split-half reliability). Độ tin cậy phân nửa được sử dụng để xác định mức độ nhất quán giữa hai nửa của cùng một phương pháp đo lường. Ví dụ, nếu bạn sử dụng một phương pháp để đo lường một phạm trù có mười chỉ số. Mười chỉ số này sẽ được chia ngẫu nhiên thành hai nửa (nhóm) mỗi nhóm năm (số lượng chỉ số mỗi nhóm có thể khác nhau nếu tổng số chỉ số là lẻ). Mỗi nhóm lúc này được xem là một công cụ dùng để đo lường một mẫu trong số khách thể (quần thể) nghiên cứu. Hệ số tương quan giữa tổng điểm của hai nhóm được xem là điểm số của độ tin cậy phân nửa. Số lượng chỉ số càng nhiều thì khả năng kết quả giống nhau của hai nhóm càng cao (bởi vì lỗi ngẫu nhiên sẽ được hạn chế khi có nhiều chỉ số được lựa chọn). Điều này có thể dẫn đến việc đánh giá quá cao độ tin cậy của những phương pháp đo lường có nhiều chỉ số.

Độ tin cậy nhất quán nội tại (Internal consistency reliability). Độ tin cậy nhất quán nội tại dùng để xác định tính nhất quán giữa hai chỉ số của cùng một phạm trù. Khi sử dụng một phương pháp để đo một phạm trù có nhiều chỉ số, mức độ những chỉ số này được đánh giá theo xu hướng tương tự nhau sẽ phản ánh mức độ nhất quán nội tại của thang đo. Độ tin cậy này có thể được xác định dựa trên độ tương quan giữa các chỉ số với chỉ số, giữa chỉ số trung bình với tổng hoặc phổ biến hơn được xác định bằng công thức Cronbach's alpha. Ví dụ, nếu như bạn có một thang 6 chỉ số, bạn sẽ có 15 cặp chỉ số khác nhau và như vậy có 15 tương quan giữa 6 chỉ số này. Tương quan trung bình giữa chỉ số với chỉ số là trung bình cộng của 15 tương quan giữa chúng. Để tính tương quan trung bình giữa chỉ số và tổng, đầu tiên bạn phải tạo ra một chỉ số "tổng" bằng cách bổ sung giá trị của cả 6 chỉ số, tính tương quan giữa chỉ số tổng với mỗi một chỉ số thành phần và cuối cùng lấy trung bình cộng của 6 tương quan vừa có được.

Cả hai phương pháp nêu trên không tính đến số lượng các chỉ số được sử dụng (6 chỉ số trong ví dụ trên). Để khắc phục điều này, phương pháp Cronbach's alpha, được thiết kế

bởi Lee Cronbach năm 1951, đã xem xét đến cả kích cỡ của thang đo (số lượng chỉ số) khi tính toán độ tin cậy của thang đo. Phương pháp Cronbach's alpha sử dụng công thức:

$$\alpha = \frac{K}{K-1} \left(1 - \frac{\sum_{i=1}^K \sigma_{Y_i}^2}{\sigma_X^2} \right)$$

trong đó K là số lượng chỉ số trong thang đo, σ_X^2 là phương sai (căn bậc hai của độ lệch chuẩn) của tổng điểm đã quan sát và $\sigma_{Y_i}^2$ là phương sai đã quan sát đối với chỉ số i. Phương pháp Cronbach's alpha được chuẩn hóa bằng cách sử dụng một công thức đơn giản hơn:

$$\alpha_{\text{standardized}} = \frac{K\bar{r}}{1 + (K-1)\bar{r}}$$

trong đó K là số lượng chỉ số trong thang đo, \bar{r} là tương quan trung bình giữa chỉ số với chỉ số, được tính bằng trung bình cộng giữa các hệ số K(K-1)/2 trong mô hình tam giác trên hoặc tam giác dưới (upper or lower triangular matrix).

Giá trị khoa học (validity)

Khái niệm **giá trị khoa học** (hay độ giá trị hoặc gọi tắt là giá trị) đề cập đến mức độ chính xác và phù hợp mà một phương pháp có được khi đánh giá phạm trù cần nghiên cứu. Ví dụ, liệu một đánh giá về “lòng từ bi” có thật sự đánh giá về lòng từ bi mà không phải là đang đánh giá một phạm trù khác như “sự đồng cảm”? Giá trị khoa học có thể được đo lường thông qua việc sử dụng các cách phương pháp lý thuyết cũng như thực nghiệm và sẽ là lý tưởng nếu kết hợp cả hai.

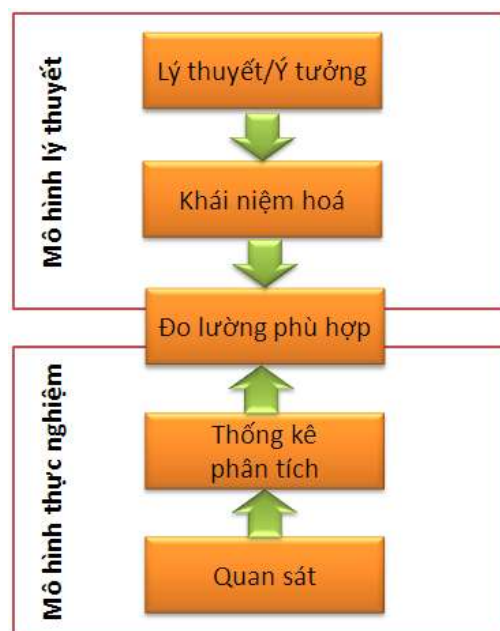
Phương pháp lý thuyết, trong xác định giá trị khoa học, nhấn mạnh mức độ xác thực của quá trình chuyển đổi (hoặc quá trình đại diện) từ một phạm trù lý thuyết thành một đánh giá thực nghiệm. Loại giá trị này được gọi là giá trị chuyển đổi (hoặc giá trị đại diện) và nó chứa đựng hai tiêu giá trị là: giá trị hình thức (face validity) và giá trị nội dung (content validity). **Giá trị chuyển đổi** (translational validity) được thẩm định bởi một hội đồng khoa học - những người sẽ đánh giá mức độ phù hợp của từng chỉ số đối với định nghĩa lý thuyết của phạm trù đó. Đồng thời, giá trị chuyển đổi còn có thể được đánh giá thông qua một kỹ thuật định tính được gọi là Q-sort.

Dựa trên các quan sát thực nghiệm, các *phương pháp thực nghiệm*, trong xác định giá trị khoa học, xác định mức độ liên quan của một công cụ đánh giá thông qua một bộ

tiêu chuẩn khách quan. Loại giá trị này được gọi là **giá trị tiêu chuẩn** (criterion-related validity), nó bao gồm bốn tiêu giá trị: giá trị tương đồng (convergent validity), giá trị phân biệt hay giá trị dị biệt (discriminant validity), giá trị đồng quy (concurrent validity) và giá trị dự báo (predictive validity).

Trong khi giá trị chuyển đổi kiểm tra liệu một đánh giá có phản ánh trung thực, đầy đủ phạm trù cơ bản hay không thì giá trị tiêu chuẩn kiểm tra liệu một đánh giá có được thực hiện theo yêu cầu mà phạm trù lý thuyết đó đã đặt ra hay không. Phương pháp thực nghiệm dùng để đánh giá giá trị khoa học dựa trên những phân tích định lượng đối với những dữ liệu đã thu thập với nhiều kỹ thuật thống kê khác nhau như phân tích tương quan, phân tích nhân tố,... Sự khác biệt giữa các phương pháp lý thuyết và phương pháp thực nghiệm trong đánh giá giá trị khoa học được minh họa trong Hình 7.2. Tuy vậy, để đảm bảo xác định đúng giá trị khoa học của các đánh giá trong nghiên cứu khoa học xã hội, cả hai cách tiếp cận này đều cần thiết và nên được kết hợp cùng với nhau.

Cũng cần phải chú ý rằng các loại giá trị khoa học khác nhau được thảo luận ở đây muốn nói tới giá trị của các *phương thức đánh giá* các phạm trù, chúng khác với *giá trị kiểm định giả thuyết* ví dụ như giá trị nhân quả, giá trị phổ quát, giá trị phạm trù và giá trị tin cậy. Các giá trị kiểm định giả thuyết sẽ được thảo luận trong một chương sau.



Hình 7.2. Hai cách tiếp cận đánh giá

Giá trị hình thức (face validity) trả lời câu hỏi liệu rằng một chỉ số có phù hợp về mặt hình thức với phạm trù cần đo lường hay không. Ví dụ như tần suất có mặt tại các buổi

giáo lễ của một người có là một chỉ số phù hợp để đo lường sự sùng đạo của người đó. Nếu ta thấy về mặt hình thức, điều này là hợp lý thì chỉ số này được xem là có giá trị hình thức. Trong khi đó nếu ta đề xuất tần suất đi tới thư viện để đọc sách của nhân viên là một chỉ số để đánh giá phạm trù hăng say làm việc của nhân viên văn phòng thì e rằng không hợp lý cho lắm, vì vậy chỉ số này có giá trị hình thức không cao.

Có một điểm thú vị đó là trong thực tế có một số chỉ số được sử dụng khá phổ biến trong các nghiên cứu về phát triển tổ chức mà khi nhìn bề ngoài có vẻ như có giá trị hình thức rất thấp. Ví dụ như năng lực tiếp thu của một tổ chức (muốn nói đến khả năng tiếp nhận, học hỏi những thông tin mới để cải thiện năng suất lao động của tổ chức), đã được sử dụng như là một chỉ số để đánh giá mức độ phát triển của tổ chức đó. Nếu nghiên cứu của bạn chứa những phạm trù rất trừu tượng hoặc những phạm trù rất khó để có cái nghĩa rõ ràng (ví dụ lòng từ bi và sự đồng cảm), bạn có thể phải nhờ đến các chuyên gia trong lĩnh vực đang nghiên cứu để đánh giá giá trị hình thức của những chỉ số dự kiến sẽ sử dụng.

Giá trị nội dung (content validity) đòi hỏi các chỉ số phải phản ánh bản chất, nội dung của phạm trù cần đo lường. Ví dụ, nếu bạn muốn đo lường phạm trù “hài lòng với dịch vụ nhà hàng” và bạn định nghĩa rằng nội dung cơ bản của dịch vụ nhà hàng bao gồm chất lượng món ăn, thái độ phục vụ của nhân viên, thời gian chờ đợi món ăn và không gian nhà hàng (ví dụ liệu trong nhà hàng có quá ồn ào hay nhiều khói), thì để có được giá trị nội dung cao, các chỉ số phải làm sao đo lường một cách chính xác mức độ khách hàng thỏa mãn với chất lượng món ăn, thái độ phục vụ, thời gian chờ đợi và không gian nhà hàng. Giá trị nội dung đòi hỏi phải lột tả chi tiết, chuẩn xác tất cả nội hàm của phạm trù nghiên cứu. Điều này tất nhiên là không dễ đối với những phạm trù phức tạp như sự tự tin, sự thông minh. Và như vậy, việc đánh giá thấu đáo giá trị nội dung của công cụ đo lường đôi khi không đạt được. Cũng giống như trong giá trị hình thức, tham vấn các chuyên gia trong lĩnh vực đang nghiên cứu là một việc làm hữu ích.

Giá trị hội tụ (hay còn gọi giá trị tương đồng - convergent validity), nói tới độ tập trung khi phản ánh phạm trù cần quan tâm của một công cụ đo lường, trong khi đó **giá trị phân biệt** hay dị biệt (discriminant validity) đề cập đến khả năng công cụ không đo lường những phạm trù không quan tâm. Thông thường thì giá trị hội tụ và giá trị phân biệt được đánh giá cùng với nhau cho một bộ phạm trù. Ví dụ khi bạn muốn kiểm chứng luận điểm: tri thức của một tổ chức (tổng hợp tri thức, kỹ năng của toàn bộ nhân viên trong một tổ chức) có liên quan tới hiệu suất làm việc của tổ chức đó, bạn cần phải đảm bảo là công cụ

đo lường tri thức của tổ chức đang thực sự đánh giá tri thức của tổ chức (giá trị hội tụ) mà không phải đánh giá hiệu suất làm việc của tổ chức đó (giá trị phân biệt).

Giá trị hội tụ có thể được xác định bằng mức độ tương tự (hoặc mức tương quan cao) trong giá trị đo lường của hai chỉ số khác nhau trong cùng một phạm trù. Giá trị phân biệt có được nhờ xác định sự khác biệt (hoặc mức tương quan thấp) trong giá trị đo lường của các chỉ số trong các phạm trù khác nhau. Trong cả hai ví dụ trên, nếu chúng ta có một công cụ với ba chỉ số để đo lường “tri thức của tổ chức” (KL) và một công cụ khác cũng với ba chỉ số để đo “hiệu suất làm việc” (PF) của tổ chức đó thì dựa trên dữ liệu có được từ mẫu đã quan sát, chúng ta có thể tính toán tương quan hai biến giữa từng cặp chỉ số trong phạm trù tri thức và phạm trù hiệu suất. Nếu mô hình tương quan này cho thấy những mối tương quan cao giữa các chỉ số của cùng một phạm trù và những mối tương quan thấp giữa các chỉ số của hai phạm trù, thì chúng ta đã cùng một lúc xác định được giá trị hội tụ và giá trị phân biệt (xem Bảng 7.3).

	KL1	KL2	KL3	PF1	PF2	PF3
KL1	1.00	0.83	0.79	0.23	0.21	0.19
KL2		1.00	0.75	0.11	0.20	0.03
KL3			1.00	0.03	-0.11	0.17
PF1				1.00	0.84	0.91
PF2					1.00	0.77
PF3						1.00

Tương quan cao giữa các chỉ số của cùng phạm trù (giá trị hội tụ)

Tương quan thấp giữa chỉ số của các phạm trù khác nhau (giá trị phân biệt)

Bảng 7.3. Phân tích tương quan hai biến để xác định giá trị hội tụ và giá trị phân biệt

Một phương pháp khác được sử dụng phổ biến hơn để xác định giá trị hội tụ và phân biệt là phương pháp *phân tích nhân tố khám phá* (exploratory factor analysis). Đây là kỹ thuật rút gọn dữ liệu bằng cách nén một bộ chỉ số lớn thành một bộ nhỏ hơn các nhân tố dựa trên cấu trúc tương quan hai biến (như đã thảo luận ở trên). Những nhân tố này cần phải phản ánh một cách trọn vẹn phạm trù lý thuyết cần đo lường. Nguyên tắc chung của việc chiết xuất nhân tố này là mỗi nhân tố được chiết xuất phải có hệ số tải nhỏ hơn 1.0.

Tùy thuộc vào khả năng tương quan của phạm trù, các nhân tố này có thể quay vòng bằng cách sử dụng kỹ thuật xoay vuông góc hoặc xoay chéo để tạo ra những hệ số tải nhân tố sử dụng để nén các chỉ số đơn lẻ thành một công cụ tích hợp. Để đạt được giá trị hội tụ cao, các chỉ số của cùng một phạm trù nên có hệ số tải 0.6 hoặc cao hơn so với một nhân tố đơn bất kỳ (hệ số tải cùng nhân tố), trong khi đó để có được giá trị phân biệt tốt, các chỉ số này nên có hệ số tải 0.3 hoặc nhỏ hơn so với tất cả các nhân tố khác (hệ số tải liên nhân tố). Mô hình này được minh họa trong bảng 7.4.

Một kỹ thuật phức tạp hơn để xác định giá trị hội tụ và phân biệt là phương pháp hỗn hợp *multi-trait multi-method* (MTMM). Phương pháp này đòi hỏi việc đo lường mỗi một phạm trù phải được thực hiện bằng hai phương pháp khác nhau (ví dụ khảo sát và quan sát cá nhân hoặc khảo sát hai nhóm tham gia nghiên cứu khác nhau như hỏi giáo viên và phụ huynh để đánh giá chất lượng học tập). Đây là phương pháp hết sức công phu và không thật sự phổ biến, vì vậy không được thảo luận kỹ ở đây.

Giá trị tiêu chuẩn cũng có thể được xác định dựa trên mức độ liên quan của một công cụ với những phạm trù khác xuất hiện ở thời điểm hiện tại hoặc trong tương lai. Phương pháp này được gọi lần lượt là giá trị đồng quy và giá trị dự báo. Giá trị dự báo (predictive validity) là mức độ thành công một công cụ có thể dự báo về những kết quả trong tương lai dựa trên lý thuyết hiện tại. Ví dụ, nếu công cụ kiểm tra học lực của học sinh trung học có khả năng phản ánh tương đối chính xác năng lực học tập của họ ở bậc đại học, thì công cụ này có giá trị dự báo cao. Việc đánh giá giá trị này đòi hỏi phải có một mạng tương tác (monological network) để chỉ ra các phạm trù liên hệ với nhau như thế nào về mặt lý thuyết.

	Nhân tố 1	Nhân tố 2
KL1	0.88	0.13
KL2	0.93	0.11
KL3	0.87	0.03
PF1	0.17	0.93
PF2	-0.03	0.85
PF3	0.07	0.78

Hệ số tải cao cho cùng nhân tố (giá trị hội tụ) Hệ số liên nhân tố thấp (giá trị phân biệt)

Bảng 7.4. Phân tích nhân tố khám phá để xác định giá trị hội tụ và phân biệt

Giá trị đồng quy (concurrent validity) đánh giá mức độ một công cụ đo lường có khả năng đánh giá một phạm trù khác đang diễn ra đồng thời. Ví dụ, nếu một công cụ dùng để đánh giá học lực của học sinh trong môn toán giải tích có một tương quan mạnh với học lực của họ trong môn đại số tuyến tính thì công cụ này có giá trị đồng quy cao. Những điểm số này có tính chất đồng quy bởi vì chúng đều phản ánh học lực của môn toán học trong cùng thời điểm. Không giống như giá trị hội tụ và phân biệt, giá trị đồng quy và dự báo thường bị bỏ qua trong các nghiên cứu thực nghiệm trong khoa học xã hội.

Lý thuyết về phương thức đánh giá

Bây giờ chúng ta đã biết có rất nhiều cách khác nhau để ước lượng độ tin cậy và giá trị khoa học của các công cụ đo lường trong khoa học xã hội, chúng ta hãy cùng tổng hợp những hiểu biết này theo lối toán học bằng cách sử dụng thuyết kiểm định cổ điển (classical test theory), cũng được gọi là thuyết điểm số thực (true score theory). Đây là một đo nghiệm tâm lý dùng để kiểm tra cách thức hoạt động của các phương thức đánh giá, chúng đánh giá cái gì và chúng không đánh giá cái gì. Lý thuyết này tổng kết rằng tất cả nghiên cứu đều có một giá trị thật T, giá trị này nó thể được nghiên cứu một cách chính xác nếu không xuất hiện các lỗi trong quá trình đánh giá. Tuy nhiên, sự hiện diện của những lỗi đánh giá E gây nên một độ lệch trong giá trị nghiên cứu so với giá trị thật như sau:

$$\begin{array}{ccccccc} X & = & T & + & E \\ \text{Giá trị nghiên cứu} & & \text{Giá trị thật} & & \text{Lỗi} \end{array}$$

Xuyên suốt các giá trị nghiên cứu, sự sai lệch của các giá trị nghiên cứu và giá trị thật cũng có thể được liên hệ bởi một phương trình tương tự:

$$\text{Độ lệch (X)} = \text{Độ lệch (T)} + \text{Độ lệch (E)}$$

Mục đích của đo nghiệm tâm lý là để ước đoán và giảm thiểu độ lệch lỗi (E), vì thế mà giá trị nghiên cứu X sẽ tiệm cận với giá trị thật T.

Có hai loại lỗi đánh giá: lỗi ngẫu nhiên và lỗi hệ thống. **Lỗi ngẫu nhiên** (random error) tạo ra bởi các yếu tố khách quan không thể phát hiện và kiểm soát được và chúng ảnh hưởng một cách ngẫu nhiên tới một số quan sát nhất định mà thôi. Ví dụ, trong thời gian tiến hành các cuộc điều tra bằng hỏi, một số người trả lời có thể không có được một tâm trạng tốt, điều này có ảnh hưởng phần nào đến câu trả lời của họ. Trong khi đó những người có tâm trạng tốt có thể có những câu trả lời tích cực hơn về một số nội dung nhất

định ví dụ như lòng tự trọng, sự thỏa mãn và niềm đam mê so với những người không có được tâm lý tốt khi trả lời bảng hỏi. Tuy nhiên, chúng ta không thể dự đoán trước được ai là người có tâm trạng tốt và độ hưng phấn hay bức xúc trong tâm trạng của họ, cũng như không thể kiểm soát được mức ảnh hưởng của những tâm trạng đó đối với nghiên cứu.

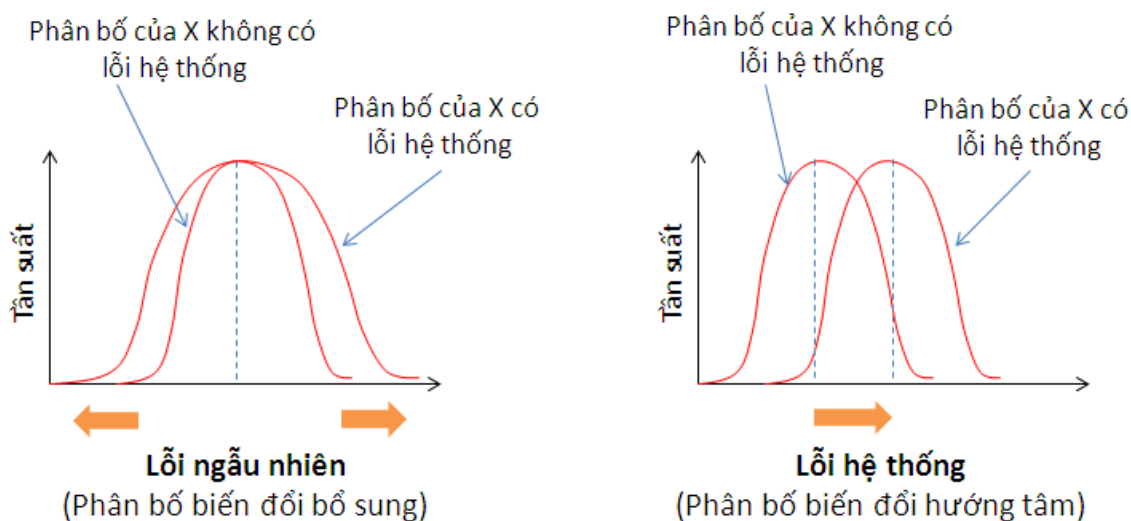
Tương tự như vậy, ở mức độ cơ quan, tổ chức. Nếu chúng ta đang đánh giá hiệu quả công tác của một đơn vị, những thay đổi về chính sách hay môi trường làm việc có thể ảnh hưởng tới một số đơn vị nhất định trong mẫu mà chúng ta đang quan sát. Vì vậy, các lỗi ngẫu nhiên được coi là những yếu tố làm nhiễu và nhìn chung được bỏ qua.

Lỗi hệ thống (systematic error) được tạo ra bởi các yếu tố ảnh hưởng có tính hệ thống đến tất cả các quan sát về một hiện tượng trên phạm vi toàn bộ mẫu nghiên cứu. Ví dụ, khi chúng ta đánh giá về hiệu quả kinh doanh của một công ty. Vào thời điểm mà cuộc khủng hoảng tài chính gần đây, đã có những ảnh hưởng bất lợi ở mức độ đáng kể hơn đối với một số công ty như các công ty trong lĩnh vực sản xuất hàng tiêu dùng và dịch vụ. Nếu như mẫu nghiên cứu chỉ chứa đựng các công ty sản xuất hàng hóa và dịch vụ, chúng ta sẽ có thể thu được kết quả nghiên cứu phản ánh sự suy giảm có hệ thống trong hiệu quả kinh doanh của tất cả các công ty trong mẫu do tác động tiêu cực và toàn diện của cuộc khủng hoảng. Không giống như các lỗi ngẫu nhiên ảnh hưởng tích cực hoặc tiêu cực hoặc không ảnh hưởng đến một số đơn vị trong mẫu, các lỗi hệ thống thường ảnh hưởng toàn diện và theo chỉ một chiều tích cực hoặc tiêu cực đối với toàn bộ đơn vị trong mẫu. Chính vì vậy, lỗi hệ thống đôi khi được xét xét là nhân tố “thiên vị” trong phương thức đánh giá và cần phải được nhìn nhận và khắc phục.

Do giá trị nghiên cứu có thể bao hàm cả lỗi ngẫu nhiên lẫn lỗi hệ thống nên phương trình giá trị thật có thể được chuyển thành:

$$X = T + E_r + E_s$$

trong đó E_r và E_s lần lượt là lỗi ngẫu nhiên và lỗi hệ thống. Ảnh hưởng về thống kê của những lỗi này là: lỗi ngẫu nhiên làm gia tăng tính biến thiên (độ lệch chuẩn) đối với sự phân bố của một đánh giá, nhưng không ảnh hưởng tới xu hướng trung tâm (giá trị trung bình); trong khi đó lỗi hệ thống tác động tới xu hướng trung tâm nhưng không ảnh hưởng tới tính biến thiên, như được minh họa trong Hình 7.5.



Hình 7.5. Ảnh hưởng của lỗi ngẫu nhiên và lỗi hệ thống.

Nghiên cứu về lỗi ngẫu nhiên và lỗi hệ thống có ý nghĩa gì đối với quy trình đánh giá? Bởi vì lỗi ngẫu nhiên làm tăng tính biến thiên của các quan sát nên chúng sẽ làm giảm độ tin cậy của phương thức đánh giá. Ngược lại, lỗi hệ thống làm thay đổi xu hướng hướng tâm của đánh giá, chúng là giảm giá trị khoa học của phương thức đánh giá. Những quan ngại về giá trị khoa học thường sâu sắc và nghiêm trọng hơn rất nhiều so với những hoài nghi về độ tin cậy, bởi vì sự đánh giá không phù hợp có thể được dùng để xem xét một phạm trù khác, không phải phạm trù chúng ta cần nghiên cứu. Vì vậy, sai lệch về giá trị khoa học gây ra những hoài nghi nghiêm trọng hơn về kết quả nghiên cứu.

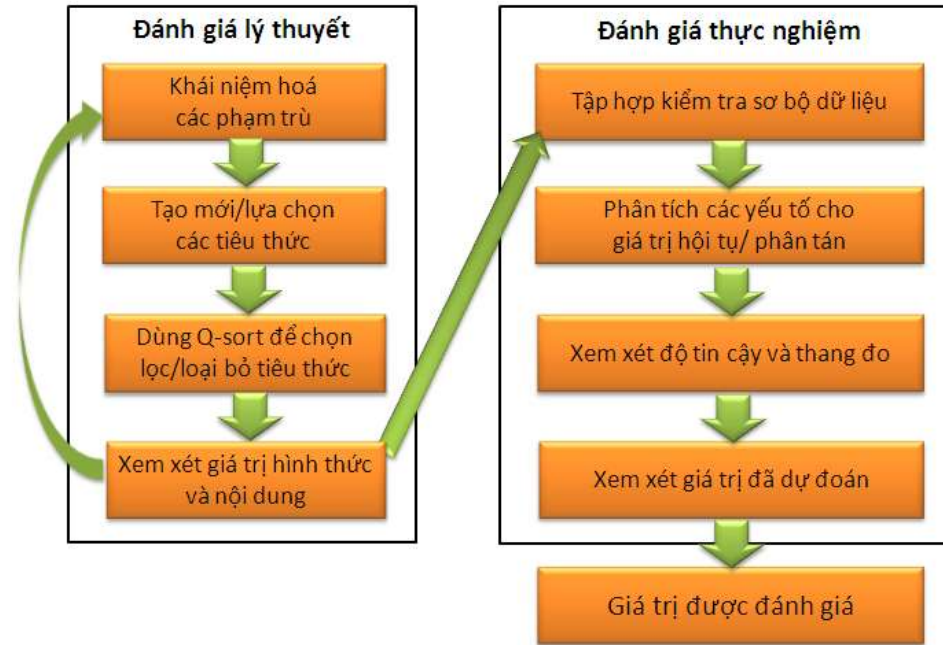
Cần chú ý rằng độ tin cậy là một tỉ lệ hoặc phân số phản ánh mối tương quan giữa giá trị thật và giá trị nghiên cứu. Vì vậy, độ tin cậy có thể được phát biểu dưới dạng:

$$\text{Độ lệch (T) / Độ lệch (X)} = \text{Độ lệch (T) / [Độ lệch (T) + Độ lệch (E)]}$$

Nếu độ lệch (T) = độ lệch (X) thì giá trị thật sẽ có tính biến thiên tương tự như giá trị nghiên cứu và như vậy độ tin tưởng là 1.0.

Phương pháp tiếp cận tổng hợp đánh giá giá trị

Một đánh giá hoàn thiện và đúng đắn về giá trị khoa học phải chứa đựng cả hướng tiếp cận đánh giá lý thuyết lẫn thực nghiệm. Như được minh họa trong Hình 7.6 dưới đây là một trình tự nhiều giai đoạn phức tạp, nó cần phải xem xét rất nhiều thang đo độ tin cậy và giá trị khoa học.



Hình 7.6. Tiếp cận tổng hợp đánh giá giá trị

Cách tiếp cận tổng hợp này bắt đầu từ đánh giá lý thuyết. Giai đoạn đầu tiên là việc khái niệm hóa những phạm trù cần nghiên cứu, nó bao gồm việc định nghĩa từng phạm trù, xác định nội hàm và ngoại diên của nó. Tiếp theo, chúng ta lựa chọn (hoặc tạo ra) những chỉ số (hay trị số) đánh giá cho mỗi phạm trù dựa trên kết quả khái niệm hóa các phạm trù như được mô tả trong Chương 5. Quá trình khảo cứu tài liệu là rất hữu ích trong việc lựa chọn các chỉ số. Mỗi một chỉ số phải được phát biểu một cách đơn giản, dễ hiểu.

Tiếp sau đó một ban thẩm định (gồm những học giả có kinh nghiệm về chuyên ngành đang nghiên cứu, về phương pháp nghiên cứu hoặc một số đại diện của những người sẽ trả lời trong nghiên cứu) sẽ kiểm định mỗi chỉ số và thực hiện phân tích Q-sort. Trong phân tích này, mỗi người trong ban thẩm định sẽ được trao một danh sách chứa tất cả phạm trù và định nghĩa của chúng cùng với các thẻ đánh giá (mỗi thẻ đánh giá một chỉ số). Sau đó mỗi vị thẩm định viên sẽ đọc lập đọc từng thẻ, kiểm tra và đánh giá văn phong, tính rành mạch và ngữ nghĩa của mỗi phát biểu (chỉ số).

Độ tin cậy nội bộ (inter-rater reliability) sẽ xác định mức độ các thẩm định viên đồng ý với các phân loại của họ. Những phát biểu được cho là không rõ ràng bởi nhiều thành viên sẽ cần phải xem xét lại, viết lại hoặc loại bỏ. Những phát biểu tốt nhất về mỗi phạm trù sẽ được lựa chọn cho các bước phân tích tiếp theo. Mỗi một phát biểu được lựa chọn sẽ được thẩm định lại để xác định giá trị hình thức và giá trị nội dung. Nếu sau giai đoạn này

có quá ít các phát biểu được lựa chọn thì phải dựa trên quá trình khái niệm hóa các phạm trù quan tâm để hình thành các phát biểu mới. Chính vì vậy, phân tích Q-sort có thể phải được thực hiện hai hoặc ba lần thì mới lựa chọn được đủ số lượng các chỉ số phản ánh toàn diện nhất, khách quan nhất các phạm trù nghiên cứu.

Tiếp theo, quy trình củng cố, cải thiện giá trị khoa học chuyển tới góc độ đánh giá qua thực nghiệm. Một công cụ nghiên cứu chứa đựng các chỉ số đã sàng lọc ở giai đoạn lý thuyết sẽ được tạo ra và áp dụng đối với một nhóm thí điểm đại diện cho khách thể nghiên cứu. Dữ liệu thu thập sẽ được lập bảng phục vụ cho phân tích tương quan hoặc phân tích thừa số với sự hỗ trợ của một phần mềm, ví dụ như phần mềm SAS hay SPSS để xác định giá trị hội tụ và giá trị phân/dị biệt. Những chỉ số không thỏa mãn các nguyên tắc về hệ số nhân tố tải lên (những hệ số tải lên của cùng nhân tố lớn hơn 0.60 và nhỏ hơn 0.30) sẽ bị loại bỏ ở bước này.

Độ tin cậy của những thang đo còn lại được xác định bằng cách sử dụng một kỹ thuật đánh giá độ nhất quán nội tại (internal consistency). Chẳng hạn phương pháp Cronbach alpha. Tiếp theo, chúng ta đánh giá khả năng dự báo của mỗi phạm trù trong một mạng tương tác được cụ thể hóa về lý thuyết bằng các kỹ thuật như phân tích hồi quy hay mô hình hóa phương trình cấu trúc. Nếu như các công cụ đánh giá phạm trù thỏa mãn hầu hết hoặc tất cả các yêu cầu đặt ra đối với giá trị khoa học và độ tin cậy được mô tả trong chương này, chúng ta có thể khẳng định rằng sau quá trình thao tác hóa, các công cụ này là cơ bản thích hợp và đúng đắn.

Giải pháp tổng hợp nhằm cải thiện giá trị khoa học được đề cập ở đây đòi hỏi khá nhiều thời gian và công sức của nhà nghiên cứu. Tuy nhiên, tiến trình nhiều giai đoạn và phức tạp này là cần thiết để đảm bảo rằng các thang đánh giá, đo lường được áp dụng trong nghiên cứu của chúng ta thỏa mãn những nguyên tắc quan trọng trong nghiên cứu khoa học. Bởi những suy luận dựa trên những thang đo sai lệch sẽ không có ý nghĩa, do đó phương thức đánh giá các phạm trù luôn là một trong những giai đoạn quan trọng nhất và phức tạp nhất của nghiên cứu thực nghiệm.

Chương 8

CHỌN MẪU

Chọn mẫu (sampling) là quá trình lựa chọn một tập hợp con (gọi là một “mẫu”) của khách thể (quần thể) nghiên cứu nhằm mục đích tiến hành nghiên cứu và đưa ra các suy luận thống kê về khách thể nghiên cứu đó. Nhìn chung, nghiên cứu khoa học xã hội là nhằm suy ra mô hình hành vi của các khách thể nghiên cứu. Do tính khả thi và hạn chế về thời gian, chi phí, chúng ta không thể nghiên cứu toàn bộ khách thể nghiên cứu, mà chúng ta phải chọn một mẫu đại diện trong số đối tượng đó để thu thập và phân tích dữ liệu. Điều quan trọng nhất trong lựa chọn mẫu là mẫu được lựa chọn phải thực sự đại diện cho khách thể nghiên cứu và như vậy các suy luận bắt nguồn từ mẫu có thể được khái quát hóa cho toàn bộ khách thể nghiên cứu. Chọn mẫu thiên vị là lý do đầu tiên giải thích cho những suy luận khác nhau và những khiếm khuyết được báo cáo trong các cuộc khảo sát và thăm dò ý kiến trước các cuộc bầu cử Tổng thống Mỹ được thực hiện bởi các tổ chức khác nhau như CNN / Gallup Poll, ABC, CBS.

Tiến trình chọn mẫu



Khách thể (Population):

Nhóm cần nghiên cứu (ví dụ: nhóm người làm việc cùng ngành trên toàn thế giới)



Khung chọn mẫu:

Một danh sách cách chọn mẫu (ví dụ: những người làm việc ở 1-2 công ty bản địa)



Chọn mẫu:

Các mẫu được chọn nghiên cứu (ví dụ: lựa chọn ngẫu nhiên những người làm việc ở mỗi công ty)

Hình 8.1. Tiến trình chọn mẫu

Chọn mẫu gồm một số công đoạn (xem hình 8.1). Đầu tiên là phải xác định rõ khách thể nghiên cứu. **Khách thể nghiên cứu** (hay quần thể nghiên cứu - population) được định nghĩa là tất cả con người hoặc sự vật (đơn vị phân tích) chứa đựng nội dung cần nghiên cứu. Đơn vị phân tích có thể là một người, một nhóm người, một tổ chức, một quốc gia, một vật thể hoặc bất kỳ một thực thể mà bạn mong muốn đưa ra những suy luận khoa học về chúng. Đôi khi khách thể nghiên cứu có thể nhìn thấy được rất rõ ràng. Ví dụ, một nhà sản xuất muốn xem liệu một dòng sản phẩm làm ra có đáp ứng yêu cầu chất lượng hay là phải làm lại hoặc loại bỏ, thì khách thể nghiên cứu là toàn bộ dòng sản phẩm đã sản xuất. Trong khi đó, chúng ta có thể thấy khó khăn hơn khi xác định khách thể nghiên cứu. Nếu bạn muốn xác định các yếu tố cơ bản tác động đến quá trình học tập của các học sinh trung học, thì khách thể nghiên cứu của bạn là học sinh trung học, giáo viên, hiệu trưởng các trường hay bậc phụ huynh? Câu trả lời đúng trong trường hợp này là học sinh trung học, bởi vì bạn quan tâm đến việc học tập của họ, không phải là việc học của các giáo viên, cha mẹ hoặc trường học.

Bước thứ hai trong quá trình chọn mẫu là xác định một **khung chọn mẫu** (sampling frame). Đây là hướng giúp nhà nghiên cứu có thể tiếp cận được với khách thể nghiên cứu. Nếu khách thể nghiên cứu của bạn là những người làm việc cùng ngành nghề thì bạn không thể tiếp cận tất cả họ trên toàn thế giới. Bởi vậy, một khung lấy mẫu khả thi hơn là danh sách những người làm việc của một hoặc hai cơ quan địa phương mà họ sẵn sàng tham gia trong nghiên cứu của bạn. Nếu khách thể nghiên cứu của bạn là công ty, thì lựa chọn trong danh sách 500 công ty hàng đầu Hoa Kỳ do Tạp chí Fortune bình chọn (Fortune 500) hay danh sách Standard & Poor (S&P) của các công ty đã đăng ký với thị trường chứng khoán New York sẽ là hướng có thể chấp nhận được.

Lưu ý rằng nếu khung chọn mẫu không hoàn toàn đại diện cho tất cả khách thể nghiên cứu thì những suy luận rút ra từ mẫu được lựa chọn cũng không thể khái quát cho toàn bộ khách thể nghiên cứu. Ví dụ, nếu khách thể nghiên cứu của bạn là tất cả nhân viên của một tổ chức (ví dụ, bạn muốn nghiên cứu lòng tự trọng, sự tự tin cá nhân của nhân viên) và khung lấy mẫu của bạn là nhân viên tại các công ty ô tô ở vùng Trung Tây nước Mỹ. Kết quả nghiên cứu phát hiện từ các nhóm này (công ty ô tô) có thể không khái quát cho lực lượng lao động toàn Mỹ chứ chưa nói đến toàn cầu. Điều này là bởi ngành công nghiệp ô tô Mỹ đang bị áp lực cạnh tranh khốc liệt trong 50 năm qua và đã có những giai đoạn cải tổ, thu hẹp quy mô lại rất nhiều - vấn đề này có thể tác động mạnh đến tinh thần và sự tự tin của nhân viên. Hơn nữa, phần lớn lực lượng lao động Mỹ đang làm việc trong

các ngành công nghiệp dịch vụ hoặc các doanh nghiệp nhỏ, không phải trong ngành công nghiệp ô tô. Do đó, một mẫu của các nhân viên ngành công nghiệp ô tô Mỹ không thật sự đại diện của lực lượng lao động của Mỹ.

Trương tự như vậy, danh sách Fortune 500 bao gồm số 500 doanh nghiệp lớn nhất của Mỹ sẽ không đại diện cho tất cả các công ty Mỹ nói chung, bởi phần đông các công ty tại Mỹ lại là các doanh nghiệp vừa và nhỏ. Do đó, nếu chọn các công ty trong danh sách Fortune 500 để làm đại diện cho toàn bộ doanh nghiệp Mỹ thì đó có thể là một khung lấy mẫu thiên vị. Ngược lại, danh sách S&P sẽ cho phép bạn để chọn trong số các công ty lớn, vừa và nhỏ, nhưng danh sách S&P lại không gồm các công ty tư nhân); cho nên đó vẫn là thiên vị.

Cũng lưu ý rằng đôi khi khách thể nghiên cứu từ đó mẫu được rút ra có thể không hoàn toàn là đối tượng mà chúng ta thực sự muốn tìm kiếm thông tin. Ví dụ, nếu một nghiên cứu muốn đánh giá mức độ thành công của cuộc vận động “bỏ thuốc lá” thì khách thể nghiên cứu là toàn bộ những người hút thuốc đã hưởng ứng và tham gia cuộc vận động này. Như vậy, rất khó để xác định rõ về toàn bộ ai là những người đã hút thuốc và đã tham gia cuộc vận động. Do đó, nhà nghiên cứu có thể lấy mẫu từ những bệnh nhân đến cơ sở y tế địa phương để điều trị cai nghiện thuốc lá. Một số người trong mẫu này có thể đã không biết hoặc đã không tham gia cuộc vận động, trong trường hợp này, khung lấy mẫu không tương ứng với khách thể nghiên cứu cần quan tâm.

Bước cuối cùng trong lấy mẫu là việc chọn các đối tượng cụ thể từ khung lấy mẫu bằng cách sử dụng một hoặc một vài phương pháp lấy mẫu nhất định. Phương pháp lấy mẫu có thể được nhóm lại thành hai loại cơ bản: lấy mẫu xác suất (ngẫu nhiên) và lấy mẫu phi xác suất. Lấy mẫu xác suất sẽ là lý tưởng nếu việc khái quát hóa kết quả là quan trọng đối với nghiên cứu của bạn, nhưng có thể có những trường hợp đặc biệt phương pháp lấy mẫu phi xác suất cũng có thể chấp nhận được. Những phương pháp này sẽ được thảo luận trong hai phần tiếp theo.

Lấy mẫu xác suất

Lấy mẫu xác suất (probability sampling) là một kỹ thuật trong đó tất cả các đơn vị trong khách thể nghiên cứu đều có cơ hội được lựa chọn trong mẫu và tỉ lệ được chọn có thể được quyết định chính xác. Vì vậy, mẫu được chọn theo phương pháp lấy mẫu xác suất sẽ mang đến những phản ánh không thiên vị về những thông số của khách thể nghiên cứu. Lấy mẫu xác suất có hai thuộc tính chung: (1) tất cả các đơn vị trong đối tượng nghiên cứu

đều có một cơ hội được lựa chọn, và (2) lấy mẫu được thực hiện ngẫu nhiên. Các kỹ thuật lấy mẫu xác suất khác nhau bao gồm:

Lấy mẫu ngẫu nhiên đơn giản (simple random sampling). Trong kỹ thuật này, tất cả các tập con của khách thể nghiên cứu (chính xác hơn, các khung lấy mẫu) đều được cho một xác suất lựa chọn bằng nhau. Xác suất lựa chọn của bất kỳ đơn vị n nào trong tổng số N đơn vị trong một khung lấy mẫu là ${}^N C_n$. Do đó, số liệu thu được từ mẫu là các đánh giá không thiên vị và không có bất kỳ đơn vị phân tích nào được nhấn mạnh. Lấy mẫu ngẫu nhiên đơn giản bao gồm đến việc chọn ngẫu nhiên người tham gia từ một khung lấy mẫu lớn.

Thông thường, để chọn mẫu, một bảng các số ngẫu nhiên hoặc một công cụ lấy số ngẫu nhiên trên máy vi tính sẽ được sử dụng. Ví dụ, nếu bạn muốn chọn 200 doanh nghiệp để khảo sát từ danh sách 1.000 công ty. Nếu danh sách 1.000 công ty này được nhập vào một bảng tính như Excel, bạn có thể sử dụng chức năng RAND[] của Excel để tạo ra các số ngẫu nhiên trong số 1.000 số được nhập. Tiếp theo, bạn sắp xếp danh sách theo thứ tự ngày càng tăng của các số ngẫu nhiên được chọn và chọn 200 số đầu tiên trong danh sách được sắp xếp. Đây là đơn giản nhất của tất cả các kỹ thuật lấy mẫu xác suất và sự đơn giản cũng là điểm mạnh của kỹ thuật này. Bởi vì khung lấy mẫu không bị phân chia hoặc phân vùng, mẫu có được là không thiên vị và những kết luận là khái quát nhất trong số tất cả các kỹ thuật lấy mẫu xác suất.

Lấy mẫu hệ thống (systematic sampling). Trong kỹ thuật này, khung lấy mẫu được xác định theo một số tiêu chí nào đó và các đối tượng được lựa chọn một cách lần lượt dựa trên khung lấy mẫu đó. Lấy mẫu hệ thống bắt đầu bằng một lựa chọn ngẫu nhiên đến đối tượng thứ k nào đó với $k = N/n$, trong đó k là tỷ lệ lấy mẫu, N là quy mô của toàn bộ khách thể nghiên cứu và n là quy mô mẫu mong muốn. Sau đó lại bắt đầu bằng một lựa chọn ngẫu nhiên. Quá trình này được tiến hành cho đến hết tất cả khách thể nghiên cứu đã dự tính. Điều quan trọng là điểm bắt đầu chọn mẫu không nhất thiết phải đối tượng đầu tiên trong danh sách, mà có thể được chọn ngẫu nhiên trong danh sách.

Trong ví dụ trước đây của chúng ta về lựa chọn 200 doanh nghiệp từ một danh sách của 1.000 doanh nghiệp. Bạn có thể sắp xếp 1.000 doanh nghiệp theo thứ tự tăng lên (hoặc giảm xuống) về quy mô của chúng (ví dụ, căn cứ vào số lượng nhân viên hay doanh thu hàng năm). Ta có tỷ lệ lấy mẫu là $k = 1.000/200 = 5$. Theo đó, chọn ngẫu nhiên một trong năm doanh nghiệp đầu tiên trong danh sách được sắp xếp và sau đó đếm đến 5 và chọn tất

cả các công ty thứ năm trong danh sách. Quá trình này sẽ đảm bảo không có sự hiện diện quá nhiều, mất cân đối của các doanh nghiệp lớn hay nhỏ trong mẫu của bạn, mà các doanh nghiệp dù lớn hay nhỏ đều có đại diện trong mẫu. Nói cách khác, về góc độ quy mô của khách thể nghiên cứu, mẫu đã chọn đại diện toàn bộ khách thể nghiên cứu.

Lấy mẫu phân tầng (stratified sampling). Trong lấy mẫu phân tầng, khung lấy mẫu được chia thành nhiều nhóm nhỏ đồng nhất và không chồng chéo (được gọi là “các tầng”) và một mẫu ngẫu nhiên đơn giản sẽ được rút ra trong mỗi nhóm. Trong ví dụ trước khi lựa chọn 200 doanh nghiệp từ một danh sách 1.000, bạn có thể bắt đầu bằng cách phân loại các công ty dựa trên quy mô của chúng: lớn (hơn 500 nhân viên), trung bình (từ 50 đến 500 nhân viên) và nhỏ (dưới 50 nhân viên). Sau đó bạn có thể lựa chọn ngẫu nhiên 67 doanh nghiệp đến từ mỗi nhóm để làm cho mẫu của bạn của 200 doanh nghiệp.

Tuy nhiên, vì trong số 1000 doanh nghiệp, số lượng các doanh nghiệp nhỏ nhiều hơn rất nhiều so với các doanh nghiệp lớn, vì vậy việc lựa chọn một con số bằng nhau (67) từ nhóm các công ty nhỏ, vừa và lớn sẽ làm cho các mẫu thu được thiếu đại diện cho toàn bộ 1.000 doanh nghiệp (tức là, thiên vị cho các công ty lớn khi tỉ lệ hiện diện của chúng trong mẫu cao hơn nhiều tỉ lệ của các doanh nghiệp nhỏ). Đây được gọi là **lấy mẫu phân tầng phi tỷ lệ** (non-proportional stratified sampling) vì tỷ lệ lấy mẫu trong mỗi nhóm không phản ánh tỷ lệ tổng thể khách thể nghiên cứu vì có trường hợp nhóm có số lượng ít hơn (ví dụ các công ty có quy mô lớn) được hiện diện trong mẫu quá nhiều.

Một kỹ thuật khác sẽ được thay thế bằng cách chọn số lượng các nhóm nhỏ tương ứng với tỉ lệ của chúng trong tổng số lượng khách thể nghiên cứu. Ví dụ, nếu có 100 doanh nghiệp lớn, 300 doanh nghiệp cỡ vừa và 600 doanh nghiệp nhỏ, thì 200 doanh nghiệp trong mẫu sẽ bao gồm 20 từ nhóm "lớn", 60 từ nhóm "trung bình" và 120 từ nhóm "nhỏ". Trong trường hợp này, sự phân phối tỷ lệ của các doanh nghiệp được giữ lại trong mẫu và vì thế kỹ thuật này được gọi là **lấy mẫu phân tầng tỷ lệ** (proportional stratified sampling).

Lấy mẫu theo cụm (cluster sampling). Nếu khách thể nghiên cứu phân tán trên một vùng địa lý rộng lớn, tiến hành lấy mẫu ngẫu nhiên đơn giản của toàn bộ khách thể nghiên cứu có thể không khả thi. Trong trường hợp này, phân chia khách thể nghiên cứu thành các “cụm” có thể là một ý tưởng hợp lý (thường là dọc theo các ranh giới địa lý), ngẫu nhiên lấy mẫu một vài cụm và đánh giá tất cả các đơn vị trong cụm đó. Ví dụ, nếu bạn muốn để lấy mẫu chính quyền các thành phố của bang New York, thay vì phải đi toàn bang để phỏng vấn các quan chức chủ chốt của các thành phố trong bang (như bạn có thể phải làm

khi lấy mẫu ngẫu nhiên đơn giản), bạn có thể nhóm các chính quyền theo đơn vị quận, lựa chọn ngẫu nhiên một bộ ba quận và sau đó chỉ phỏng vấn các quan chức chủ chốt trong ba quận đó. Tuy có sự khác biệt giữa các cụm lấy mẫu, độ sai lệch ước tính của chọn mẫu theo cụm là cao hơn so với lấy mẫu ngẫu nhiên; do đó, tính khái quát cho toàn thể khách thể nghiên cứu kém hơn so với kết quả thu được từ kỹ thuật lấy mẫu ngẫu nhiên đơn giản.

Lấy mẫu so sánh cặp (matched-pairs sampling). Đôi khi, các nhà nghiên cứu muốn so sánh hai phân nhóm trong khách thể nghiên cứu dựa trên một tiêu chí cụ thể nào đó. Ví dụ, tại sao một số công ty liên tục gặt hái nhiều lợi nhuận hơn so với các công ty khác? Để tiến hành một nghiên cứu như vậy, bạn sẽ phải phân loại khung lấy mẫu thành các công ty có lợi nhuận cao và các công ty có lợi nhuận thấp dựa trên tỷ suất lợi nhuận gộp, thu nhập trên mỗi cổ phiếu hoặc một số tiêu chí đo lường khác về lợi nhuận. Sau đó bằng kỹ thuật lấy mẫu ngẫu nhiên đơn giản bạn sẽ chọn một công ty trong mỗi nhóm, dựa trên sự tương xứng về quy mô, lĩnh vực kinh doanh, và/hoặc các tiêu chí phù hợp khác. Bây giờ, bạn có hai mẫu xuất hiện của các công ty lợi nhuận cao và lợi nhuận thấp mà bạn có thể nghiên cứu chi tiết hơn. Như vậy, kỹ thuật lấy mẫu so sánh cặp thường là một cách lý tưởng để nghiên cứu sự khác biệt lưỡng cực giữa các phân nhóm khác nhau trong một khách thể nghiên cứu nhất định.

Lấy mẫu nhiều giai đoạn (multi-stage sampling). Tất cả các kỹ thuật chọn mẫu xác suất được mô tả bên trên là các kỹ thuật lấy mẫu một giai đoạn. Tùy thuộc vào nhu cầu lấy mẫu của bạn, bạn có thể kết hợp các kỹ thuật một giai đoạn để tiến hành lấy mẫu nhiều giai đoạn. Ví dụ, bạn có thể phân tầng một danh sách của các doanh nghiệp dựa trên quy mô doanh nghiệp và sau đó tiến hành lấy mẫu hệ thống trong mỗi tầng. Đây là một sự kết hợp hai kỹ thuật: lấy mẫu phân tầng và lấy mẫu hệ thống. Tương tự như vậy, bạn có thể bắt đầu với một nhóm quận trong tiểu bang New York: trong mỗi quận, chọn một mẫu ngẫu nhiên đơn giản các trường học; trong mỗi trường, chọn một mẫu ngẫu nhiên đơn giản của các lớp và trong mỗi lớp, chọn một mẫu ngẫu nhiên đơn giản các sinh viên để nghiên cứu. Trong trường hợp này, bạn có một quá trình lấy mẫu bao gồm bốn giai đoạn của lấy mẫu theo cụm và lấy mẫu ngẫu nhiên đơn giản.

Chọn mẫu phi xác suất

Chọn mẫu phi xác suất (non-probability sampling) là một kỹ thuật lấy mẫu, trong đó một số đơn vị của khách thể nghiên cứu không có cơ hội được chọn hoặc khả năng được chọn không thể được xác định chính xác. Thông thường, các đơn vị được lựa chọn dựa trên

một số tiêu chí không ngẫu nhiên, chẳng hạn như hạn ngạch hay sự thuận tiện. Bởi vì lựa chọn là không ngẫu nhiên, chọn mẫu không xác suất khiến cho việc đánh giá các lỗi lấy mẫu trở nên khó khăn và có thể phải chấp nhận sự thiên vị trong lấy mẫu. Do đó, thông tin từ một mẫu không thể được khái quát hóa cho toàn bộ khách thể nghiên cứu. Các loại kỹ thuật chọn mẫu phi xác suất bao gồm:

Chọn mẫu thuận tiện (convenience sampling). Cũng được gọi là lấy mẫu tình cờ hoặc lấy mẫu cơ hội, đây là một kỹ thuật mà trong đó một mẫu được rút ra từ một phần của khách thể nghiên cứu mà chúng gần gũi, có sẵn hoặc thuận tiện cho nhà nghiên cứu. Ví dụ, nếu bạn đứng bên ngoài một trung tâm mua sắm và trao các bảng câu hỏi cho khách hàng hoặc phỏng vấn họ khi họ bước vào trong trung tâm mua sắm. Những mẫu mà bạn nhận được dạng này là một mẫu thuận tiện. Đây là một mẫu phi xác suất bởi vì bạn bỏ qua những khách hàng tại các trung tâm mua sắm khác. Các ý kiến nhận được từ mẫu này có thể phản ánh những đặc điểm đặc trưng của trung tâm mua sắm này, chẳng hạn như phương châm kinh doanh của trung tâm (ví dụ, chỉ hướng đến những khách hàng trung lưu) hoặc vị trí của trung tâm (ví dụ, gần một trường đại học sẽ thu hút nhiều sinh viên), do đó không thể đại diện cho ý kiến của toàn thể người mua sắm. Do đó, khả năng khái quát hóa của những quan sát như vậy sẽ hạn chế. Ví dụ khác về lấy mẫu thuận tiện là lấy mẫu những sinh viên trong một lớp học hoặc lấy mẫu các bệnh nhân đến một phòng khám y tế nhất định. Đây là loại lấy mẫu hữu ích trong giai đoạn thí điểm các nghiên cứu với mục tiêu thử nghiệm các công cụ nghiên cứu hoặc kiểm định các chỉ số đánh giá, chứ không nên sử dụng để đưa ra các kết luận khái quát.

Chọn mẫu định mức (quota sampling). Trong kỹ thuật này, khách thể nghiên cứu được phân chia thành nhiều nhóm nhỏ (giống như trong lấy mẫu phân tầng) và sau đó một số lượng không ngẫu nhiên sẽ được chọn từ mỗi nhóm nhỏ để đáp ứng một định mức được xác định trước. Trong **chọn mẫu định mức tỷ lệ** (proportional quota sampling), tỷ lệ số người được chọn trong mỗi nhóm phải phù hợp với cấu trúc của khách thể nghiên cứu. Ví dụ, nếu dân số Mỹ bao gồm 70% người da trắng, 15% người Mỹ gốc Tây Ban Nha, 13% người Mỹ gốc Phi; bạn muốn hiểu quan điểm bầu cử của họ trong một mẫu của 98 người, bạn có thể đứng bên ngoài một trung tâm mua sắm và đề nghị mọi người bỏ phiếu quan điểm của họ. Bạn sẽ phải dừng việc lấy phiếu của người gốc Tây Ban Nha khi bạn đã có 15 phản hồi (hoặc dừng lấy phiếu người gốc Phi khi bạn đã có 13 phản hồi), để các thành phần dân tộc trong mẫu phù hợp với tỉ lệ dân số Mỹ nói chung.

Chọn mẫu định mức phi tỷ lệ (non-proportional quota sampling) không bị hạn chế bởi yêu cầu phải đạt được một tỷ lệ đại diện nào đó, nhưng có thể phải đạt được số lượng tối thiểu trong mỗi nhóm. Trong trường hợp này, bạn có thể quyết định có 50 người trả lời từ mỗi nhóm trong ba nhóm dân tộc (người da trắng, người Mỹ gốc Tây Ban Nha và người Mỹ gốc Phi) và bạn dừng lại khi đạt đến định mức của bạn cho mỗi phân nhóm. Không một loại lấy mẫu định mức nào là có thể đại diện của toàn dân số Mỹ, bởi vì tùy thuộc vào nơi nghiên cứu được thực hiện (ví dụ tại một trung tâm mua sắm ở New York hay Kansas) kết quả có thể hoàn toàn khác nhau. Kỹ thuật lấy mẫu định mức có tính đại diện thấp, nhưng có thể hữu ích bởi nó cho phép nắm bắt được ý kiến của các nhóm nhỏ.

Chọn mẫu chuyên gia (expert sampling). Đây là một kỹ thuật mà người được lựa chọn một cách không ngẫu nhiên dựa trên chuyên môn của họ về lĩnh vực đang nghiên cứu. Ví dụ, để hiểu về tác động của một chính sách mới của chính phủ như Đạo luật Sarbanes-Oxley^(*), bạn có thể lấy mẫu một nhóm kế toán, những người quen thuộc với lĩnh vực này. Ưu điểm của phương pháp này là ở chỗ các chuyên gia có hiểu biết sâu rộng hơn về vấn đề nghiên cứu, nên ý kiến từ mẫu của các chuyên gia đáng tin cậy hơn mẫu bao gồm cả những người không có chuyên môn. Tuy vậy, những phát hiện từ mẫu sẽ không khái quát hóa cho toàn bộ quần thể nói chung.

Chọn mẫu mở rộng dần (chọn mẫu quả bóng tuyết - snowball sampling). Trong chọn mẫu quả cầu tuyết, bạn bắt đầu bằng cách nhận diện một vài người phù hợp với các tiêu chí để tiếp cận cho nghiên cứu của bạn và sau đó đề nghị họ giới thiệu những người khác mà họ biết sẽ đáp ứng các tiêu chí lựa chọn. Ví dụ nếu bạn muốn khảo sát quản trị mạng máy tính, ban đầu bạn chỉ biết một hoặc hai chuyên gia quản trị mạng, sau khi thu nhận thông tin từ họ, bạn có thể đề nghị họ giới thiệu những người khác cũng làm quản trị mạng. Mặc dù phương pháp này hầu như không dẫn đến mẫu có tính đại diện, nhưng đôi khi nó có thể là cách duy nhất để tiếp cận với những khách thể nghiên cứu khó tiếp cận hoặc khi không có khung lấy mẫu nào sẵn có.

Thống kê mẫu (Statistics of Sampling)

Trong phần trước, chúng ta đã đưa ra các thuật ngữ như tham số quần thể (thông số về toàn bộ khách thể nghiên cứu), thống kê mẫu và sai số chọn mẫu. Trong phần này,

^(*) Đạo luật Sarbanes-Oxley còn được biết với tên Đạo luật Sarbox, là một trong những luật căn bản của nghề kế toán, kiểm toán, được ban hành tại Hoa Kỳ năm 2002. Mục tiêu chính của Đạo luật này nhằm bảo vệ lợi ích của các nhà đầu tư vào các công ty đại chúng bằng cách buộc các công ty này phải cải thiện sự đảm bảo và tin tưởng vào các báo cáo, các thông tin tài chính công khai (theo Wikipedia - N.D).

chúng ta cố gắng làm rõ nghĩa và mối quan hệ của các thuật ngữ này.

Khi đo lường một quan sát nào đó từ một đơn vị mẫu, chẳng hạn như câu trả lời của một cá nhân cho một câu hỏi trong bảng hỏi được thiết kế theo thang đo Likert, thì kết quả quan sát đó được gọi là một phản hồi hay một câu trả lời (xem Hình 8.2). Nói cách khác, một câu trả lời (response) là một giá trị đo lường đưa ra bởi một đơn vị trong mẫu đã chọn. Các câu trả lời cho cùng một câu hỏi (hay một chỉ số, một quan sát) từ những người trả lời khác nhau có thể được đồ thị hóa thông qua đường **phân bố tần suất** (frequency distribution), nó phản ánh tần suất xuất hiện của câu trả lời. Mỗi phiếu trả lời cung cấp cho bạn những câu trả lời khác nhau cho từng mục trong phiếu khảo sát. Các ý kiến, quan điểm của các phiếu khảo sát khác nhau cho cùng vấn đề hoặc đánh giá có thể biểu diễn bằng đồ thị tần suất phân bố các sự kiện xảy ra. Đối với số lượng lớn các câu trả lời trong một mẫu khảo sát, đồ thị phân bố tần suất theo đường cong giống hình quả chuông và được coi là phân bố chuẩn (normal distribution). Đồ thị phân bố có thể dùng để đánh giá đặc tính chung của một mẫu, như giá trị trung bình của mẫu (trung bình cộng của tất cả các kết quả trả lời trong một mẫu) hoặc độ lệch chuẩn (mức độ biến thiên hoặc giao thoa của các câu trả lời trong một mẫu). Toàn bộ thông số về mẫu như vậy được gọi là *thống kê mẫu* – sample statistics (một “thống kê” là một giá trị được đánh giá từ các dữ liệu thu được).

Các quần thể cũng có cùng giá trị trung bình và độ lệch chuẩn nếu mẫu được chọn là toàn bộ quần thể (hay toàn bộ khách thể nghiên cứu). Tuy nhiên, vì chọn mẫu là toàn bộ quần thể là không thực tế nên các đặc tính của cả quần thể không bao giờ có thể xác định được thấu đáo. Những đặc tính này được gọi là các *thông số quần thể* - population parameters (không thể gọi chúng là “thống kê” được vì chúng không thể xác định được thông qua các kỹ thuật phân tích dữ liệu thống kê). Thông tin từ thống kê mẫu có thể khác với các thông số quần thể khi mẫu không thật sự đại diện cho quần thể. Sự khác biệt giữa thống kê mẫu và các thông số quần thể được gọi là **sai số chọn mẫu** (sampling error). Về mặt lý thuyết, nếu chúng ta tăng dần kích cỡ mẫu đến mức tiệm cận cả quần thể thì sẽ giảm được sai số chọn mẫu. Như vậy, thống kê mẫu ngày càng tiệm cận với thông số quần thể tương ứng.

Nếu một mẫu được chọn thực sự đại diện cho quần thể thì thông tin trong các thống kê mẫu sẽ đồng nhất với các thông số quần thể tương ứng. Làm cách nào để biết được thống kê mẫu đồng nhất hoặc ít nhất là phản ánh sát các thông số quần thể ở mức độ có thể chấp nhận được? Trước hết, chúng ta cần hiểu khái niệm **phân bố chọn mẫu** (sampling distribution).

Tên tiêu chí	ATT1	ATT2	ATT3	INT1	INT2	INT3	FREQ	HRS
	6	6	6	6	6	6	2	2
	1	1	1	1	1	1	3	10
Tất cả câu trả lời của một phiếu trả lời	6	5	4	2	3	3	2	2
	6	6	6	5	5	5	2	4
	6	6	6	6	5	4	2	2.5
	7	7	7	7	7	6	2	5
	5	4	5	5	4	5	2	1
	6	6	6	6	7	7	3	10
Tập hợp câu trả lời cho một tiêu chí của tất cả các phiếu trả lời tạo thành phân bố tần suất	7	6	6	5	5	4	2	1
	4	4	4	4	4	4	1	0
	5	5	5	4	4	4	2	1
	4	4	4	4	4	4	2	5
	6	6	6	5	5	5	2	4
	-1	-1	-1	-1	-1	-1	-1	-1
	-1	-1	-1	-1	-1	-1	-1	-1
	-1	-1	-1	-1	-1	-1	-1	-1
Trung bình (hoặc SD) của tập hợp này là một thống kê mẫu	6	6	6	6	6	6	2	15
	6	6	5	5	5	4	2	2
	1	1	1	1	1	1	1	0
	6	6	6	6	6	6	2	6
	6	6	5	6	6	5	2	3
	5	5	5	2	2	2	2	7
	4	5	6	1	3	2	2	3
	6	7	7	6	7	6	2	3

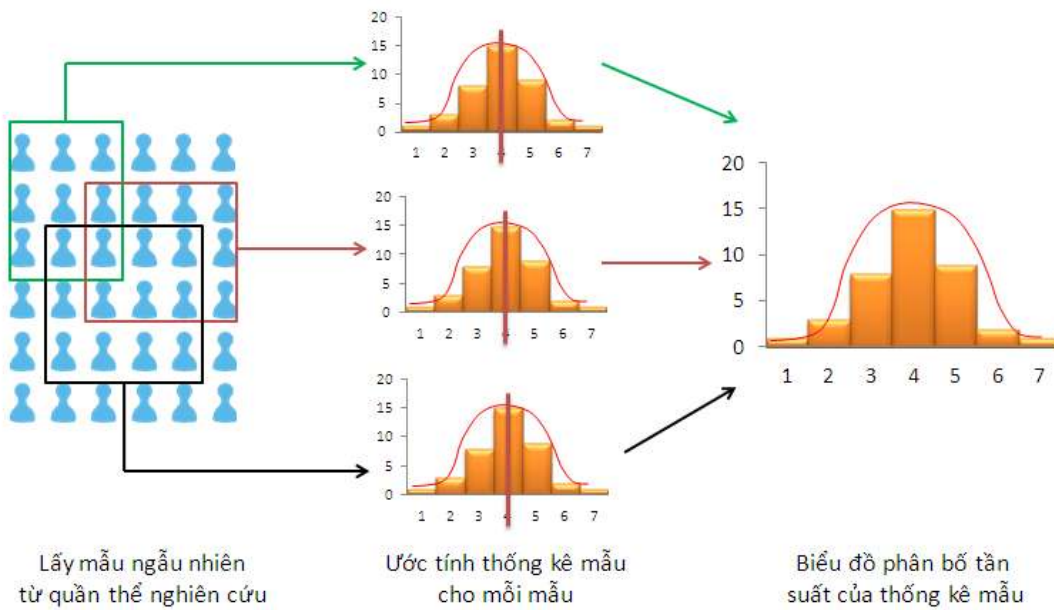
Hình 8.2. Thống kê mẫu

Hãy tưởng tượng là bạn lấy 3 mẫu ngẫu nhiên khác nhau của một quần thể nghiên cứu như trong Hình 8.3. Đối với mỗi mẫu, bạn nhận được các giá trị thống kê mẫu như giá trị trung bình, độ lệch chuẩn. Nếu mỗi mẫu ngẫu nhiên thực sự đại diện cho quần thể, thì khi đó giá trị trung bình của 3 mẫu sẽ đồng nhất (và bằng với thông số quần thể) và độ biến thiên trong giá trị trung bình của mẫu bằng 0.

Nhưng điều này là vô cùng hiếm, bởi vì mỗi một mẫu ngẫu nhiên có thể rút ra từ những phần khác nhau trong quần thể. Do đó, giá trị trung bình của các mẫu có sự khác biệt nhất định. Tuy nhiên, có thể sử dụng giá trị trung bình của 3 mẫu và vẽ đồ thị tần suất cho các giá trị trung bình. Nếu số lượng mẫu tăng từ 3 lên 10 hoặc lên tới 100, biểu đồ tần suất trở thành một phân bố chọn mẫu. Như vậy, một phân bố chọn mẫu là một phân bố tần suất của thống kê mẫu (như giá trị trung bình mẫu) từ một tập hợp các mẫu, trong khi đó phân bố tần suất là tham chiếu chung cho sự phân bố của một câu trả lời (các đánh giá) từ một mẫu riêng biệt.

Cũng giống như phân bố tần suất, phân bố chọn mẫu nhằm mục đích có được các thống kê mẫu với phần lớn có trị số tương đương giá trị trung bình (đây là giá trị được cho là phản ánh các thông số quần thể) và một phần nhỏ hơn có trị số khác xa giá trị trung

binh. Với lượng mẫu vô cùng lớn, việc phân bố mẫu sẽ tiệm cận với phân bố thường. Độ biến thiên hoặc giao thoa của một thống kê mẫu trong một phân bố chọn mẫu (ví dụ: độ lệch chuẩn của thống kê chọn mẫu) được gọi là **sai số chuẩn** (standard error). Trái lại, thuật ngữ độ lệch chuẩn chỉ dành riêng cho độ biến thiên của câu trả lời từ một mẫu riêng biệt.



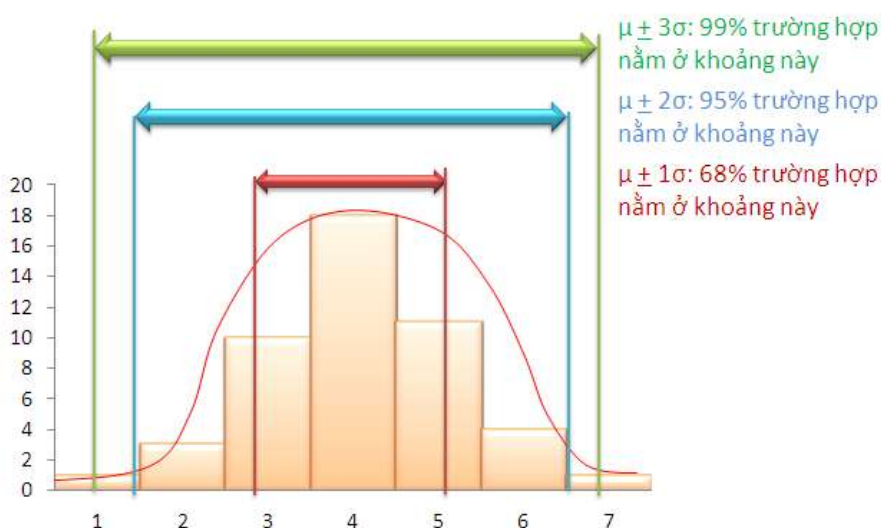
Hình 8.3. Sự phân bố chọn mẫu

Giá trị trung bình của thống kê mẫu trong một phân bố chọn mẫu được coi như là một ước lượng cho các thông số chưa biết của quần thể. Căn cứ vào dải phân bố chọn mẫu (ví dụ: căn cứ vào sai số chuẩn), chúng ta có thể xác định được các khoảng tin cậy cho các dự đoán về thông số quần thể. Khoảng tin cậy (confidence intervals) là một ước đoán xác suất mà một thông số quần thể nằm trong một khoảng cụ thể của các giá trị thống kê mẫu. Mọi phân bố thường sẽ có xu hướng đạt tới chuẩn chung 68-95-99% (xem Hình 8.4 dưới đây). Trong đó, hơn 68% trường hợp phân bố nằm trong phạm vi độ lệch chuẩn của giá trị trung bình ($\mu \pm 1\sigma$); hơn 95% trường hợp phân bố nằm trong phạm vi hai lần độ lệch chuẩn giá của trị trung bình ($\mu \pm 2\sigma$) và hơn 99% trường hợp phân bố nằm trong phạm vi ba lần độ lệch chuẩn giá trị trung bình ($\mu \pm 3\sigma$). Bởi lẽ, phân bố chọn mẫu với số lượng mẫu không hạn định sẽ tiệm cận mức phân bố thường, giống như quy tắc 68-95-99. Có thể tóm lược như sau:

- (Thống kê mẫu + một sai số chuẩn) tương ứng 68% khoảng tin cậy cho thông số quần thể.

- (Thống kê mẫu + hai sai số chuẩn) tương ứng 95% khoảng tin cậy cho thông số quần thể.
- (Thống kê mẫu + ba sai số chuẩn) tương ứng 99% khoảng tin cậy cho thông số quần thể.

Một mẫu là “bất đối xứng” (hay thiên vị - biased, không đại diện cho quần thể) nếu phân bố chọn mẫu của nó không đạt được hoặc không tuân theo quy tắc 68-95-99%. Cũng cần lưu ý rằng trong hầu hết các phân tích hồi quy, khi chúng ta xem xét ý nghĩa của hệ số hồi quy với $p < 0.05$, chúng ta phải xem xét liệu thống kê chọn mẫu (hệ số hồi quy) có dự đoán thông số quần thể tương ứng (mức ảnh hưởng thực) với khoảng tin cậy 95% hay không. Rất thú vị, chuẩn “sáu xích-ma” (six sigma) có khả năng phát hiện các sai sót ngoài khoảng in cậy 99% hoặc sáu độ lệch tiêu chuẩn (độ lệch tiêu chuẩn được ký hiệu bằng chữ cái Hy Lạp xích-ma), biểu thị ý nghĩa kiểm định ở mức $p < 0.01$.



Hình 8.4. Quy tắc khoảng tin cậy tỷ lệ 68-95-99 phần trăm

Chương 9

NGHIÊN CỨU KHẢO SÁT

Nghiên cứu khảo sát là phương pháp liên quan đến việc sử dụng các bảng câu hỏi hoặc các cuộc phỏng vấn được chuẩn hoá từ trước để thu thập một cách có hệ thống dữ liệu về người, sở thích, suy nghĩ và hành vi của họ. Mặc dù các cuộc điều tra dân số được tiến hành rất sớm từ thời Ai Cập cổ đại, song khảo sát với tư cách là một phương pháp nghiên cứu khoa học chính thức chỉ bắt đầu từ những năm 1930 - 1940 và người đi tiên phong là nhà xã hội học Paul Lazarsfeld khi ông xem xét các tác động của các đài phát thanh đối với việc hình thành các quan điểm chính trị tại Hoa Kỳ. Từ đó, phương pháp này trở thành kỹ thuật rất phổ biến cho các nghiên cứu định lượng trong khoa học xã hội.

Phương pháp khảo sát có thể được sử dụng cho các nghiên cứu mô tả, thăm dò hoặc giải thích. Nó thích hợp nhất cho các nghiên cứu đơn vị phân tích là những cá nhân con người. Mặc dù các đơn vị phân tích khác, chẳng hạn như các nhóm, tổ chức, hay những cặp đôi (ví dụ hai tổ chức, người mua và người bán,...), cũng sử dụng phương pháp khảo sát để nghiên cứu, nhưng những nghiên cứu như vậy thường sử dụng một người cụ thể từ mỗi đơn vị như là một “người phát ngôn” hoặc là “người được ủy nhiệm” của tổ chức đó. Các cuộc điều tra như vậy có thể bị thiên vị đơn nếu người phát ngôn được lựa chọn không có đủ kiến thức hoặc có ý kiến thiên vị về hiện tượng quan tâm. Ví dụ, người giám đốc điều hành có thể không biết đầy đủ về quan điểm của nhân viên công ty, do đó có thể là người cung cấp thông tin sai cho các nghiên cứu về động lực hoặc sự tự tin của nhân viên.

Nghiên cứu khảo sát có những điểm mạnh vốn có so với các phương pháp nghiên cứu khác. *Trước tiên*, các cuộc điều tra là một phương tiện tuyệt vời để đánh giá hàng loạt những dữ liệu không quan sát được như sở thích của người dân (ví dụ, định hướng chính trị), đặc điểm (ví dụ sự tự tin), thái độ (ví dụ đối với những người nhập cư), niềm tin (ví dụ về một dự luật mới), hành vi (ví dụ hút thuốc lá hoặc uống rượu) hoặc thông tin thực tế về họ (ví dụ thu nhập). *Thứ hai*, nghiên cứu khảo sát cũng rất lý tưởng để thu thập dữ liệu từ xa về một khách thể nghiên cứu mà nó quá lớn để trực tiếp quan sát. Với một khu vực rộng lớn, chẳng hạn như một quốc gia, dữ liệu có thể thu được bằng cách sử dụng thư, thư điện tử hoặc điện thoại từ một mẫu nhỏ nhưng được chọn lựa cẩn thận để đảm bảo rằng đối tượng nghiên cứu được đại diện đầy đủ trong mẫu đó.

Thứ ba, do tính chất kín đáo và khả năng đáp ứng yêu cầu thuận tiện cho người trả lời, các bảng hỏi khảo sát được ưa thích bởi nhiều người. *Thứ tư*, các cuộc phỏng vấn có thể là cách duy nhất để tiếp cận một số nhóm khách thể nghiên cứu như những người vô gia cư hoặc những người nhập cư bất hợp pháp mà không có hướng chọn mẫu sẵn có. *Thứ năm*, một số nghiên cứu khảo sát cho phép phát hiện các vấn đề nhỏ kể cả trong các cuộc điều tra quy mô lớn với nhiều biến số và cũng có thể cho phép đưa ra phân tích so sánh của các nhóm khách thể nghiên cứu (ví dụ, những so sánh ngay trong một nhóm và giữa các nhóm với nhau). *Thứ sáu*, nghiên cứu khảo sát sẽ tiết kiệm hơn về thời gian, công sức và chi phí hơn so với hầu hết các phương pháp khác như nghiên cứu thực nghiệm và nghiên cứu trường hợp. Tuy nhiên, nghiên cứu khảo sát cũng có một số nhược điểm đặc thù. Nó có thể gây ra một số vấn đề quan trọng như tỉ lệ phản hồi thấp, thiên vị trong lấy mẫu của nhà nghiên cứu, thiên vị trong trả lời của người tham gia,... Tất cả sẽ được thảo luận trong phần cuối của chương này.

Tùy thuộc vào phương thức thu thập dữ liệu, nghiên cứu khảo sát có thể được chia thành hai loại lớn: các cuộc điều tra bảng hỏi (có thể cả qua thư, thực hiện nhóm hoặc các cuộc khảo sát trực tuyến), các cuộc điều tra phỏng vấn (có thể là cá nhân, điện thoại hoặc thảo luận nhóm nhỏ). Bảng câu hỏi là công cụ được người trả lời hoàn thành bằng ngôn ngữ viết, trong khi đó các cuộc phỏng vấn được hoàn thành do người tham gia phỏng vấn dựa trên phản ứng bằng lời nói. Như bàn luận dưới đây, mỗi loại có những điểm mạnh và điểm yếu riêng về các khía cạnh như chi phí, phạm vi khách thể nghiên cứu và tính linh hoạt nhà nghiên cứu.

Điều tra bảng hỏi

Bảng hỏi (questionnaire) do Sir Francis Galton (1822-1911) phát kiến, là một công cụ nghiên cứu được tiêu chuẩn hóa, bao gồm một tập hợp các câu hỏi nhằm mục đích nắm bắt phản hồi từ người trả lời. Câu hỏi có thể được cấu trúc sẵn (structured) hoặc không được cấu trúc sẵn (unstructured). Câu hỏi không được cấu trúc sẵn (gọi tắt là câu hỏi cấu trúc) sẽ để cho người trả lời để cung cấp thông tin bằng ngôn từ của họ, trong khi các câu hỏi đã cấu trúc sẵn (gọi tắt là câu hỏi phi cấu trúc) yêu cầu người trả lời lựa chọn một câu trả lời từ một bộ có sẵn. Thông tin thu được từ các câu hỏi riêng lẻ (các chỉ số) trong bảng hỏi có thể được tổng hợp thành một thanh phức hợp hoặc thành các chỉ số so sánh, là cơ sở cho các phân tích thống kê sau này. Câu hỏi nên được thiết kế sao cho người trả lời có thể đọc, hiểu và trả lời một cách có ý nghĩa. Chính vì thế phương pháp khảo sát có thể không

phù hợp hoặc không thực tế cho một số nhóm nhân khẩu nhất định, chẳng hạn như trẻ em hoặc những người không biết đọc, viết.

Hầu hết các cuộc điều tra bằng hỏi có xu hướng là **khảo sát qua thư** (self-administered mail surveys), lúc này các bảng hỏi tương tự được gửi đến một số lượng lớn người dân. Những người quan tâm đến việc trả lời bảng hỏi có thể hoàn thành việc trả lời khi thấy thuận tiện và gửi lại trong phong bì bưu chính đã trả phí trước. Khảo sát qua thư thuận lợi ở chỗ chúng không phô trương và không quá tốn kém để thực hiện, bởi vì ở hầu hết các quốc gia, chi phí bưu chính không quá tốn kém. Tuy nhiên, tỷ lệ trả lời từ các cuộc điều tra qua thư có thể khá thấp vì hầu hết mọi người có thể bỏ qua các phiếu khảo sát. Họ cũng có thể trì hoãn lâu (vài tháng) trong việc hoàn thành trả lời và gửi trở lại (hoặc chỉ đơn giản là họ có thể làm mất nó). Do đó, các nhà nghiên cứu phải liên tục theo dõi phản ứng của họ và gửi lời nhắc nhở lặp đi lặp lại (hai hoặc ba nhắc nhở khoảng thời gian từ một đến 1,5 tháng là lý tưởng). Điều tra bằng hỏi cũng không thật sự phù hợp cho các vấn đề cần làm rõ chi tiết về người trả lời hoặc những vấn đề đòi hỏi phải có những văn bản trả lời chi tiết. Những thiết kế theo chiều dọc có thể được sử dụng để khảo sát cùng số người được hỏi ở những thời điểm khác nhau, nhưng tỷ lệ trả lời có xu hướng giảm đáng kể ở cuộc điều tra tiếp theo.

Loại thứ hai của cuộc khảo sát là **bảng hỏi thực hiện theo nhóm** (group - administered questionnaire). Những người trả lời được chọn trong mẫu tập hợp tại một nơi phù hợp về thời gian và không gian, từng người được đề nghị hoàn thành các câu hỏi điều tra tại đó. Họ trả lời một cách độc lập mà không cần tương tác với nhau. Hình thức này là thuận tiện cho các nhà nghiên cứu và tỷ lệ trả lời cao sẽ được đảm bảo. Nếu người trả lời không hiểu bất kỳ câu hỏi cụ thể nào, họ có thể yêu cầu nhà nghiên cứu làm rõ. Trong nhiều tổ chức, việc tập hợp một nhóm các nhân viên trong một phòng hội nghị, phòng ăn trưa là tương đối dễ dàng, đặc biệt nếu cuộc khảo sát được chấp thuận bởi người đứng đầu tổ chức.

Một hình thức khảo sát khác được sử dụng gần đây là các cuộc khảo sát trực tuyến. Các khảo sát này được thực hiện qua Internet bằng cách sử dụng các hình thức tương tác khác nhau. Người trả lời có thể nhận được một thư điện tử yêu cầu tham gia trong cuộc khảo sát với một liên kết (đường link) đến một trang web trực tuyến, nơi cuộc điều tra có thể được hoàn thành. Ngoài ra, cuộc khảo sát có thể được gửi vào trong e-mail, sau khi hoàn thành có thể gửi lại thông qua trả lời e-mail. Những cuộc điều tra như thế này rất không tốn kém, kết quả được ghi lại ngay lập tức trong một cơ sở dữ liệu trực tuyến và

cuộc điều tra có thể dễ dàng sửa đổi nếu cần thiết. Tuy nhiên, nếu trang web khảo sát không có mật khẩu bảo vệ, các câu trả lời có thể dễ dàng bị xâm nhập. Hơn nữa, thiên vị trong chọn mẫu có thể là một vấn đề lớn vì các cuộc khảo sát không thể tiếp cận tới những người không sử dụng máy tính truy cập Internet, như các nhóm người nghèo, người cao tuổi và dân tộc thiểu số. Các bảng hỏi trực tuyến thường gửi tới nhóm nhân khẩu trẻ có nhiều thời gian trực tuyến và có kỹ năng vi tính để hoàn thành cuộc điều tra như vậy. Việc ước tính tỷ lệ trả lời có thể có vấn đề, nếu đường link tới trang khảo sát được đăng trên một diễn đàn hoặc bảng thông báo mà không email trực tiếp đến người cung cấp thông tin. Vì lý do này mà các nhà nghiên cứu có thể sử dụng một lúc hai phương tiện truyền thông (ví dụ, khảo sát qua thư và khảo sát trực tuyến), cách này cho phép trả lời lựa chọn phương pháp trả lời mà họ ưa thích.

Xây dựng một bảng câu hỏi khảo sát là một nghệ thuật. Nhiều công việc cần được xúc tiến như nội dung của câu hỏi, từ ngữ, định dạng và trình tự; tất cả đều có thể có mang đến những tác động quan trọng đối với câu trả lời.

Hình thức trả lời (Response format). Câu hỏi khảo sát có thể được cấu trúc hoặc không được cấu trúc. Các câu trả lời đối với những câu hỏi cấu trúc có thể thiết kế bằng một trong các định dạng sau đây:

- **Câu trả lời phân đôi** (dichotomous response): người trả lời được yêu cầu chọn một trong hai sự lựa chọn, chẳng hạn như đúng - sai, có - không hoặc đồng ý - không đồng ý. Một ví dụ của câu hỏi này là: Bạn có nghĩ rằng án tử hình là hợp lý trong một số trường hợp (khoanh tròn): có - không.

- **Câu trả lời xưng danh** (nominal response): người trả lời chọn một trong số nhiều lựa chọn (nhiều hơn hai lựa chọn) và những sự lựa chọn này không được sắp xếp theo thứ tự, chẳng hạn: bạn làm trong lĩnh vực gì: dịch vụ sản xuất / công nghiệp / bán lẻ / giáo dục / y tế / du lịch & khách sạn.

- **Câu trả lời thứ tự** (ordinal response): người trả lời chọn một trong số nhiều lựa chọn (nhiều hơn hai lựa chọn) và những lựa chọn này được sắp xếp theo thứ tự, chẳng hạn như: trình độ giáo dục cao nhất của bạn là gì: trung học phổ thông/ cao đẳng / đại học/ sau đại học.

- **Câu trả lời nhiều mức độ** (interval - level response): người trả lời lựa chọn các câu trả lời trong thang Likert 5 điểm hoặc 7 điểm hoặc thang Guttman, thang đối nghĩa. Các loại thang này đã thảo luận trong chương trước.

- *Câu trả lời liên tiếp* (continuous response), người trả lời điền vào một giá trị liên tục (trong thang tỷ lệ) với điểm số không có ý nghĩa (meaningful zero point), chẳng hạn như tuổi tác hay nhiệm kỳ của họ trong một công ty. Những câu trả lời này thường là thể loại điền vào khoảng trống.

Nội dung và cách diễn đạt câu hỏi. Các câu trả lời thu được trong nghiên cứu khảo sát chịu ảnh hưởng rất nhiều bởi các loại câu hỏi. Các câu hỏi mơ hồ hoặc sắp xếp không hợp lý có thể sẽ dẫn đến những câu trả lời vô nghĩa với rất ít giá trị. Dillman (1978) khuyến cáo một số quy tắc để tạo ra các câu hỏi khảo sát tốt. Nhà nghiên cứu nên xem xét kỹ lưỡng mỗi một câu hỏi trong cuộc khảo sát theo các phương diện sau đây:

- *Câu hỏi rõ ràng và dễ hiểu.* Các câu hỏi khảo sát nên được nêu lên với một ngôn ngữ đơn giản về mặt ngữ pháp tốt nhất nên sử dụng thể chủ động, tránh thể bị động, không có các từ phức tạp hay biệt ngữ gây bối rối cho người trả lời. Tất cả các câu hỏi trong bản câu hỏi nên được diễn đạt một cách tương tự để làm cho người trả lời dễ dàng để đọc và hiểu chúng. Ngoại lệ duy nhất là khi đối tượng khảo sát của bạn là nhóm chuyên môn, chẳng hạn như bác sĩ, luật sư và các nhà nghiên cứu, những người sử dụng thuật ngữ khoa học đó trong môi trường làm việc hàng ngày của họ.

- *Câu hỏi về vấn đề tiêu cực.* Những câu hỏi đụng chạm khía cạnh tiêu cực, nhạy cảm, chẳng hạn như “theo bạn chính quyền địa phương có nên tăng thuế không”, sẽ làm khó người trả lời và có thể dẫn đến những phản hồi không chính xác. Tất cả câu hỏi như vậy nên tránh.

- *Câu hỏi mơ hồ.* Câu hỏi khảo sát không nên có những cụm từ hoặc phát biểu có thể được giải thích khác nhau từ người trả lời khác nhau. Ví dụ, nếu bạn hỏi người trả lời “thu nhập hàng năm là bao nhiêu”. Đó là câu hỏi mơ hồ, không rõ ràng vì chưa rõ bạn chỉ đề cập tiền lương (theo tháng hay theo tuần) hoặc cổ tức, hay tiền cho thuê, hay các thu nhập khác; bạn chỉ đề cập thu nhập cá nhân, hay cả thu nhập gia đình (bao gồm cả tiền lương vợ chồng) hoặc cá nhân và thu nhập doanh nghiệp? Giải thích khác nhau của những người trả lời khác nhau sẽ dẫn đến những câu trả lời không thể so sánh và sau này không thể tổng hợp một cách chính xác.

- *Câu hỏi thiên vị hoặc câu hỏi định hướng.* Thiên vị đề cập đến bất kỳ phần nào của một câu hỏi mà khuyến khích người đọc trả lời một hướng nào đó. Kenneth Rasinky (1989) đã kiểm tra một số nghiên cứu về thái độ của người dân đối với chi tiêu của chính phủ. Ông nhận thấy rằng người trả lời có xu hướng đề nghị chính phủ hỗ trợ mạnh mẽ hơn

trong vấn đề “trợ giúp người nghèo” và mà ít quan tâm hơn tới vấn đề “phúc lợi xã hội”, mặc dù cả hai từ này có ý nghĩa tương tự. Cũng trong nghiên cứu này, số người ủng hộ “ngăn chặn tội phạm đang gia tăng” nhiều hơn những người ủng hộ cho “cải cách thực thi pháp luật”, “giải quyết các vấn đề của các thành phố lớn” nhiều hơn là cho “hỗ trợ các thành phố lớn” và “đổi phó với nghiện ma túy” nhiều hơn cho “cai nghiện ma túy”. Ngôn ngữ thiên vị có thể bóp méo kết quả trả lời. Thường rất khó để dự đoán các từ ngữ thiên vị, nhưng với mức độ tốt nhất có thể, câu hỏi khảo sát nên được nghiên cứu kỹ lưỡng để tránh những từ ngữ này.

- *Câu hỏi đa nghĩa.* Đó là những câu hỏi có thể có nhiều câu trả lời. Ví dụ, “bạn có hài lòng với phần cứng và phần mềm được cung cấp cho công việc của bạn?” Trong ví dụ này, làm thế nào để người được hỏi có thể trả lời đơn nếu anh ấy / cô ấy hài lòng với phần cứng nhưng không hài lòng với các phần mềm hoặc ngược lại? Tốt nhất là nên tách câu hỏi như vậy thành hai câu riêng biệt: (1) “Bạn có hài lòng với phần cứng được cung cấp cho công việc của bạn”, và (2) “bạn có hài lòng với phần mềm được cung cấp cho công việc của bạn”.

- *Câu hỏi quá chung chung.* Đôi khi, những câu hỏi quá chung chung có thể không truyền đạt chính xác tới nhận thức của người trả lời. Nếu bạn hỏi một người nào đó mức độ họ thích một cuốn sách và cung cấp một thang trả lời với các mức độ khác nhau, từ “hoàn toàn không thích” đến “vô cùng thích”. Trong trường hợp người đó lựa chọn “vô cùng thích” thì bạn hiểu “vô cùng thích” cụ thể thế nào? Thay bằng những câu hỏi chung chung này, hãy hỏi câu hỏi cụ thể hơn về hành vi, chẳng hạn như: bạn sẽ giới thiệu cuốn sách này cho những người khác hoặc bạn có ý định đọc những cuốn sách khác của cùng tác giả...? Tương tự như vậy, thay vì hỏi công ty của bạn lớn như thế nào (câu hỏi này có thể được giải thích khác nhau bởi những người trả lời), bạn có thể hỏi có bao nhiêu người làm việc cho công ty của bạn và (hoặc) doanh thu hàng năm của công ty của bạn là bao nhiêu. Những thông tin định lượng cụ thể như vậy sẽ mang lại những thông số có ý nghĩa hơn để đánh giá quy mô doanh nghiệp.

- *Câu hỏi quá chi tiết.* Tránh những câu hỏi chi tiết một cách không cần thiết mà không có mục đích nghiên cứu cụ thể. Ví dụ, nghiên cứu của bạn chỉ cần quan tâm đến số lượng trẻ em trong gia đình, nhưng bạn hỏi đến cả tuổi của mỗi đứa trẻ trong một hộ gia đình? Tuy nhiên, nếu không chắc chắn, câu hỏi quá chi tiết vẫn còn tốt hơn câu hỏi quá tổng quát.

- *Câu hỏi giả định.* Nếu bạn hỏi, “lợi ích của việc cắt giảm thuế là gì”, trước đó bạn có thể đã giả định rằng người trả lời sẽ thấy cắt giảm thuế là có lợi. Nhưng nhiều người không nghĩ vậy, bởi vì cắt giảm thuế thường dẫn đến kinh phí chi cho giáo dục ít đi, chất lượng các dịch vụ công như cảnh sát, xe cứu thương, cứu hoả giảm xuống.

- *Câu hỏi tưởng tượng.* Một câu hỏi phổ biến trong các chương trình trò chơi truyền hình “nếu bạn kiếm được một triệu đô la từ chương trình này, bạn sẽ có kế hoạch để chi tiêu nó như thế nào?”. Hầu hết người được hỏi chưa bao giờ có được số tiền như vậy và chưa bao giờ nghĩ về nó (hầu hết đều không biết rằng tại Hoa Kỳ sau thuế, họ sẽ chỉ nhận được khoảng \$640.000 và trong nhiều trường hợp, số tiền đó được trả trong khoảng thời hạn 20 năm, do đó, giá trị thực tại rông của nó là thậm chí ít hơn). Do đó, câu trả lời của họ có xu hướng được khá ngẫu hứng, chẳng hạn như một tour du lịch trên khắp thế giới, mua một nhà hàng hoặc quầy bar, chi tiêu cho giáo dục, tiết kiệm cho hưu trí, giúp các bậc cha mẹ hoặc con cái hoặc có một đám cưới xa hoa. Câu hỏi tưởng tượng sẽ nhận được câu trả lời tưởng tượng, không thể được sử dụng trong suy luận khoa học.

- *Câu hỏi giả định người trả lời có đủ thông tin cần thiết để trả lời đúng.* Thông thường, chúng ta giả định rằng người tham gia nghiên cứu đã có các thông tin cần thiết để trả lời câu hỏi, nhưng trong thực tế lại không đúng như vậy. Trong trường hợp này, các câu trả lời thu được có thể không chính xác, vì người trả lời thiếu kiến thức về các câu hỏi nêu ra. Ví dụ, không nên hỏi giám đốc điều hành về tất cả chi tiết hoạt động hàng ngày trong công ty của họ vì họ có thể không để ý và nắm bắt tất cả các chi tiết nhỏ; hoặc không nên hỏi học sinh trung học “Bạn có nghĩ rằng chính phủ Mỹ đã có những hành động thích hợp trong xử lý vụ khủng hoảng Vịnh Con lợn (The Bay of Pigs) không?”

Thứ tự các câu hỏi. Nhìn chung, các câu hỏi nên được sắp xếp theo dòng chảy hợp lý từ đầu đến cuối. Để đạt được tỷ lệ trả lời cao nhất, các câu hỏi nên được bố trí từ ít nhạy cảm nhất đến nhạy cảm nhất, từ thực tế và hành vi tới thái độ, từ tổng quát đến cụ thể hơn. Một số quy tắc chung cho việc sắp xếp câu hỏi:

- Bắt đầu với câu hỏi dễ. Lựa chọn tốt là về nhân khẩu học (tuổi, giới tính, trình độ học vấn) cho các cuộc điều tra cá nhân và thông tin tổ chức (số nhân viên, doanh thu hàng năm, lĩnh vực) cho các cuộc điều tra tổ chức.
- Đừng bao giờ bắt đầu với một câu hỏi mở.
- Nếu theo một chuỗi các sự kiện lịch sử, theo một thứ tự thời gian từ xa nhất đến gần đây nhất.

- Hãy hỏi về một chủ đề tại một thời điểm. Khi chuyển chủ đề, sử dụng câu chuyển đổi, chẳng hạn như "Phần tiếp theo sẽ xem xét ý kiến của bạn về..."

- Sử dụng các câu hỏi lọc hoặc dự phòng khi cần thiết. Chẳng hạn như: "Nếu bạn trả lời 'có' trong câu hỏi 5, xin vui lòng chuyển tới Phần 2. Nếu bạn trả lời 'không' xin vui lòng chuyển tới Phần 3".

Các quy tắc vàng khác. Hãy đặt bản thân mình vào hoàn cảnh của người trả lời. Hãy chú ý và đánh giá cao thời gian của người trả lời, quan tâm, tin tưởng và chú ý bảo mật thông tin cá nhân của họ. Luôn luôn thực hiện các chiến lược sau đây cho tất cả các nghiên cứu khảo sát:

- Thời gian của mọi người là rất giá trị. Hãy tôn trọng thời gian của họ. Giữ cuộc khảo sát của bạn càng ngắn càng tốt và chỉ đề cập những gì thật sự cần thiết. Dù có quan trọng như thế nào, người trả lời không thích dành hơn 10-15 phút cho bất kỳ cuộc điều tra nào. Khảo sát càng dài thì tỉ lệ trả lời càng thấp.

- Luôn luôn đảm bảo sự bảo mật cho những người trả lời khi sử dụng thông tin của họ trong nghiên cứu khoa học và trong báo cáo kết quả.

- Đối với các cuộc điều tra tổ chức, phải đảm bảo rằng bạn sẽ gửi cho những người trả lời bản sao của kết quả cuối cùng và chắc chắn rằng bạn giữ lời hứa.

- Cảm ơn sự tham gia người trả lời trong nghiên cứu của bạn.

Cuối cùng, luôn luôn kiểm tra trước các câu hỏi của bạn, ít nhất là bằng cách sử dụng một mẫu thuận tiện, trước khi phát chúng tới tay người trả lời. Việc kiểm định trước như vậy có thể phát hiện ra sự mơ hồ, thiếu rõ ràng hoặc những thiên vị trong từ ngữ cần phải được loại bỏ.

Điều tra phỏng vấn

Phỏng vấn là một hình thức thu thập dữ liệu có tính cá nhân hơn so với các bảng hỏi và được thực hiện bởi người phỏng vấn có kỹ năng với hình thức thu thập thông tin tương tự như trong khảo sát bảng hỏi (ví dụ như một bộ câu hỏi được tiêu chuẩn hóa). Tuy nhiên, không giống như một bảng hỏi, kịch bản phỏng vấn có thể chứa đựng những chỉ dẫn đặc biệt cho người phỏng vấn mà người trả lời không được chứng kiến. Phỏng vấn cũng sẽ tạo cơ hội cho người phỏng vấn ghi lại những nhận xét và ý kiến cá nhân. Ngoài ra, không giống như các cuộc điều tra qua thư, điều tra phỏng vấn có cơ hội để làm rõ thêm những vấn đề do người trả lời nêu ra hoặc người hỏi đưa thêm những câu hỏi thăm dò hoặc gợi ý.

Tuy nhiên, các cuộc phỏng vấn thường tốn thời gian và công sức. Người phỏng vấn cần có những kỹ năng đặc biệt để thành công. Người phỏng vấn cũng được coi là một phần trong công cụ đánh giá và phải chủ động cố gắng không thiên vị với người được phỏng vấn.

Hình thức điển hình nhất của phỏng vấn là phỏng vấn cá nhân hay **phỏng vấn trực diện** (face-to-face interview), khi đó người phỏng vấn làm việc trực tiếp với người tham gia để đặt câu hỏi và ghi lại câu trả lời của họ. Các cuộc phỏng vấn cá nhân có thể được tiến hành tại nhà của người tham gia hoặc ở văn phòng. Cách tiếp cận này có thể được ưa chuộng bởi một số người trả lời, trong khi những người khác có thể cảm thấy không thoải mái trong việc cho phép một người lạ vào nhà của họ. Tuy nhiên, người phỏng vấn có kỹ năng có thể thuyết phục người trả lời hợp tác, cải thiện đáng kể tỉ lệ trả lời.

Một biến thể của phỏng vấn cá nhân là **phỏng vấn nhóm**, còn được gọi là thảo luận nhóm nhỏ (focus group). Trong kỹ thuật này, một nhóm nhỏ người tham dự (thường là 6-10 người) được phỏng vấn cùng nhau ở một địa điểm chung. Người phỏng vấn có trách nhiệm chủ yếu là người hỗ trợ, dẫn dắt cuộc thảo luận và đảm bảo rằng mỗi người tham gia đều có cơ hội trả lời. Thảo luận nhóm nhỏ cho phép kiểm tra sâu sắc hơn những vấn đề phức tạp so với các hình thức khác của nghiên cứu khảo sát, bởi vì khi người ta nghe thấy người khác thảo luận, họ có thể có những phản ứng hay ý tưởng mà họ không suy nghĩ trước đó. Tuy nhiên, thảo luận nhóm nhỏ có thể bị lấn át bởi một số cá nhân với tính cách áp đặt, trong khi một số cá nhân có thể phải miễn cưỡng nói lên ý kiến của mình trước mặt các đồng nghiệp hoặc cấp trên của họ, đặc biệt là khi đối phó với một vấn đề nhạy cảm như làm việc thiếu trách nhiệm nhân viên hoặc những vấn đề chính trị của đơn vị. Do kích thước mẫu nhỏ, thảo luận nhóm nhỏ thường được sử dụng trong nghiên cứu thăm dò hơn là nghiên cứu mô tả hoặc giải thích.

Loại thứ ba của điều tra phỏng vấn là **phỏng vấn qua điện thoại**. Trong kỹ thuật này, người phỏng vấn liên hệ với người trả lời tiềm năng qua điện thoại, thường dựa trên một sự lựa chọn ngẫu nhiên từ một thư mục điện thoại, để hỏi các câu hỏi khảo sát. Một cách tiếp cận gần đây sử dụng công nghệ tiên tiến hơn là phỏng vấn qua điện thoại với sự trợ giúp của máy tính (CATI). Hình thức này đang được sử dụng ngày một phổ biến bởi các viện nghiên cứu, cơ quan chính phủ và các nhà khảo sát thương mại, nơi mà người phỏng vấn là người điều hành điện thoại, được hỗ trợ xuyên suốt quá trình phỏng vấn bởi một chương trình máy tính với các hướng dẫn và câu hỏi hiển thị trên màn hình máy tính. Hệ thống này cũng lựa chọn ngẫu nhiên người trả lời bằng cách sử dụng một kỹ thuật quay số ngẫu nhiên và ghi lại các câu trả lời bởi sử dụng công nghệ thu âm. Bởi vì người trả lời

luôn giữ liên lạc bằng điện thoại, chúng ta có thể thu được tỉ lệ trả lời cao hơn. Kỹ thuật này không phải là lý tưởng cho các khu vực nông thôn, mật độ điện thoại thấp và cũng không thể được sử dụng để giao tiếp thông tin không có âm thanh như đồ họa hoặc triển lãm sản phẩm.

Vai trò của người phỏng vấn

Người phỏng vấn có vai trò rất quan trọng và đa dạng trong quá trình phỏng vấn. Để phỏng vấn, người phỏng vấn phải thực hiện các công việc sau đây:

- *Chuẩn bị cho cuộc phỏng vấn:* Bởi người phỏng vấn đi đầu trong các nỗ lực thu thập dữ liệu và chất lượng của dữ liệu thu thập phụ thuộc rất nhiều vào quá trình đào tạo người phỏng vấn. Người phỏng vấn phải được trang bị kiến thức về quá trình phỏng vấn và các phương pháp khảo sát, thám nhuần mục đích nghiên cứu, hiểu rõ cách thức lưu trữ và sử dụng thông tin người trả lời cung cấp và hiểu được các thiên vị có thể phát sinh khi phỏng vấn. Anh / cô ấy cũng nên luyện tập trước khi chính thức phỏng vấn.

- *Xác định địa chỉ và tranh thủ sự hợp tác của người trả lời:* Nhất là cuộc điều tra cá nhân thực hiện trong nhà, người phỏng vấn phải xác định địa chỉ cụ thể, tìm hiểu thời gian hợp lý mà người trả lời có thể hợp tác. Nên xem họ như người bán hàng, họ bán các ý tưởng khi tham gia nghiên cứu.

- *Động viên người trả lời:* người trả lời rất để ý tới tâm trạng của người phỏng vấn. Nếu người phỏng vấn là lơ là hoặc không tỏ ra nhiệt tình, người trả lời sẽ không có hứng thú cung cấp thông tin hữu ích. Người phỏng vấn phải chứng minh sự nhiệt tình, nhấn mạnh và truyền đạt tầm quan trọng của đề tài nghiên cứu tới người trả lời và chú ý tới tâm lý của người trả lời trong suốt cuộc phỏng vấn.

- *Làm rõ tất cả những nhầm lẫn hoặc mối lo âu:* người phỏng vấn cần phải có khả năng tinh tế phát hiện điều lo ngại hay khó chịu của người trả lời và cố gắng giải quyết nó để bầu không khí giao tiếp cởi mở. Ngoài ra, họ có thể hỏi thêm các câu thăm dò khi cần thiết ngay cả khi những câu hỏi như vậy là không có trong kịch bản.

- *Đánh giá chất lượng của các câu trả lời:* người phỏng vấn là người có được điều kiện thuận lợi nhất để đánh giá chất lượng của thông tin đã thu thập và có thể phải chủ động tìm kiếm các câu trả lời bổ sung bằng cách sử dụng quan sát cử chỉ cá nhân, ngôn ngữ cơ thể khi phù hợp.

Tiến hành phỏng vấn

Trước khi phỏng vấn, người phỏng vấn cần chuẩn bị các giấy tờ và dụng cụ cần thiết cho buổi phỏng vấn, bao gồm thư giới thiệu cơ quan nghiên cứu, nhà tài trợ, bản sao đầy đủ của các công cụ khảo sát, ảnh nhận diện, số điện thoại của người trả lời để gọi và xác minh sự xác thực của người phỏng vấn. Để bắt đầu cuộc phỏng vấn, anh / cô ấy nên nói chuyện với một ngữ điệu khẩn trương và tự tin. Chẳng hạn như: “Tôi muốn nhờ bạn một vài phút để phỏng vấn cho một nghiên cứu rất quan trọng”, thay vì “Tôi có thể phỏng vấn anh được không?”. Người phỏng vấn nên tự giới thiệu bản thân, các thông tin cá nhân hiện tại, giải thích mục đích của nghiên cứu trong 1-2 câu; đảm bảo tính bảo mật và sự tự nguyện tham gia của người trả lời, tất cả không đến một phút. Không “đao to búa lớn”, không dùng biệt ngữ, không đi sâu chi tiết trừ khi có yêu cầu cụ thể. Nếu người phỏng vấn muốn ghi âm cuộc phỏng vấn, cần có sự cho phép rõ ràng của người trả lời trước khi thực hiện. Ngay cả khi cuộc phỏng vấn được ghi âm, người phỏng vấn vẫn cần phải ghi chép các vấn đề quan trọng.

Trong suốt cuộc phỏng vấn, người phỏng vấn cần thực hiện theo các câu hỏi trong kịch bản và đặt câu hỏi chính xác như trong văn bản; không thay đổi từ ngữ để làm cho câu hỏi có vẻ thân thiện hơn. Tương tự, cũng không nên thay đổi thứ tự các câu hỏi hoặc bỏ qua bất kỳ câu hỏi nào. Tất cả vấn đề cần hỏi được thảo luận trong quá trình diễn tập trước khi đến buổi phỏng vấn thực tế. Người phỏng vấn không nên ngắt lời người trả lời. Nếu người được hỏi đưa ra một câu trả lời sơ sài, người phỏng vấn phải gợi ý để người trả lời đưa ra một phản ứng chi tiết hơn. Một số kỹ thuật gợi ý hữu ích là:

- Im lặng. Tạm dừng và chờ đợi (mà không đi sâu hoặc sang câu hỏi tiếp theo) để gợi ý cho người trả lời rằng người phỏng vấn đang chờ đợi phản ứng chi tiết hơn.
- Công khai khuyến khích. Thỉnh thoảng có thể khuyến khích người trả lời đi vào chi tiết hơn. Tuy nhiên, người phỏng vấn không được thể hiện là ủng hộ hoặc phản đối ý kiến của người trả lời.
- Đề nghị làm rõ. Chẳng hạn như “bạn có thể làm rõ thêm vấn đề đó” hoặc “bạn vừa nói về một kinh nghiệm mà bạn đã có ở trường trung học. Bạn có thể cho tôi biết thêm về điều đó”.
- Phản ánh. Người phỏng vấn có thể thử phương pháp của nhà tâm lý trị liệu bằng việc nhắc lại những gì người trả lời đã nói. Ví dụ: “Tôi vừa nghe rằng bạn đã có những trải

kinh nghiệm rất đau buồn”, tạm dừng và sau đó và chờ đợi cho người trả lời làm rõ.

Sau khi phỏng vấn hoàn thành, người phỏng vấn nên cảm ơn người trả lời vì đã dành thời gian cho cuộc phỏng vấn; hẹn với họ thời điểm sẽ có kết quả nghiên cứu và không kết thúc vội vàng. Ngay khi kết thúc, họ nên viết lại tất cả thông tin quan trọng để sau này có thể giúp giải thích ý kiến của người trả lời tốt hơn.

Những thiên vị trong nghiên cứu khảo sát

Mặc có rất nhiều những điểm mạnh và lợi thế, nghiên cứu khảo sát thường không tránh khỏi những thiên vị có hệ thống có thể làm giảm giá trị một số kết luận thu được từ các cuộc điều tra. Năm thiên vị đó là: tỉ lệ phản hồi, quá trình chọn mẫu, tính khách quan của câu trả lời, khả năng hồi tưởng của người trả lời và phương pháp phổ biến.

Tỉ lệ phản hồi. Nhìn chung các nghiên cứu khảo sát luôn bị đánh giá về tỉ lệ phản hồi thấp. Khoảng 15-20% là tỉ lệ điển hình trong một cuộc điều tra qua thư, cho dù đã có đến hai hoặc ba thư nhắc nhở. Nếu đa số khách thể nghiên cứu không trả lời thì câu hỏi đặt ra sẽ là vì sao lại họ là thờ ơ với cuộc khảo sát và sau đó là những hoài nghi về tính hợp lệ và giá trị của các kết quả của nghiên cứu. Ví dụ, các khách hàng không hài lòng có thể lên tiếng nhiều hơn so với khách hàng đã hài lòng và nhiều khả năng họ sẽ tham gia nhiều hơn vào các cuộc điều tra so với khách hàng đã hài lòng. Do đó, thành phần trong mẫu có thể có một tỷ lệ cao hơn của khách hàng không hài lòng. Trong trường hợp này, không chỉ kết quả thiếu khái quát, mà còn có thể không đúng thực tế. Một số chiến lược có thể được sử dụng để cải thiện tỉ lệ phản hồi:

- *Thông báo trước:* Một lá thư ngắn gửi trước cho những người trả lời để thu hút sự tham gia của họ trong cuộc khảo sát sắp tới có thể giúp họ chuẩn bị trước và cải thiện thái độ của họ đối với cuộc điều tra. Bức thư cần nêu rõ mục đích và tầm quan trọng của nghiên cứu, cách thức thu thập dữ liệu (ví dụ, thông qua một cuộc gọi điện thoại, một mẫu khảo sát trong thư,...) và đánh giá cao sự hợp tác của họ. Một biên thể của kỹ thuật này có thể đề nghị họ gửi lại một tấm bưu thiếp đã trả cước để thấy liệu họ có sẵn sàng tham gia vào nghiên cứu hay không.

- *Nội dung phù hợp:* Nếu một cuộc điều tra xem xét các vấn đề liên quan hay quan trọng với người trả lời, thì khả năng trả lời của họ có thể cao hơn so với những khảo sát không quan trọng với họ.

- *Câu hỏi thân thiện:* Câu hỏi điều tra ngắn hơn thường có tỷ lệ trả lời cao hơn so với

bảng câu hỏi dài. Hơn nữa, các câu hỏi rõ ràng, không gây khó chịu và dễ dàng để trả lời cũng có thể nhận được tỷ lệ trả lời cao hơn.

- *Sự hỗ trợ*: Đối với khảo sát các tổ chức, nếu người đứng đầu đứng ra xác nhận tầm quan trọng của nghiên cứu cho tổ chức, thì nhân viên sẽ nhiệt tình trả lời. Việc xác nhận như vậy có thể dưới hình thức một lá thư giới thiệu để có thể cải thiện độ tin cậy của nhà nghiên cứu trong mắt của người trả lời.

- *Nhắc nhở thường xuyên*: Các nhắc nhở liên tục có thể thuyết phục một số người từ không trả lời thành trả lời, ngay cả khi câu trả lời của họ có thể đến trễ.

- *Đào tạo người phỏng vấn*: Tỷ lệ phản hồi trong các cuộc phỏng vấn có thể được cải thiện với những người phỏng vấn có tay nghề cao, được đào tạo về cách thức phỏng vấn, sử dụng các kỹ thuật quay số trên máy vi tính để xác định những người trả lời tiềm năng, lên lịch tiếp cận lại những người chưa gặp.

- *Khích lệ bằng vật chất*: Tỷ lệ đáp ứng, ít nhất là một số người tham gia nhất định, có thể gia tăng cùng với việc sử dụng các khích lệ như tiền mặt hoặc thẻ quà tặng, vé xổ số, rút thăm trúng thưởng, phiếu giảm giá, đóng góp từ thiện,...

- *Khích lệ tinh thần*: Nhất là với các doanh nghiệp, họ sẽ tích cực tham gia nếu tặng họ những khích lệ phi vật chất hơn là các khuyến khích tiền tệ. Ví dụ tặng họ một báo cáo so sánh phản hồi của các cá nhân đối với doanh nghiệp sau khi tổng hợp tất cả các câu trả lời trong cuộc khảo sát.

- *Bảo mật và sự riêng tư*: Cuối cùng, đảm bảo rằng những thông tin cá nhân mà người trả lời cung cấp sẽ không bị hé lộ tới bất kỳ bên thứ ba nào. Đây là cách có thể giúp cải thiện tỷ lệ phản hồi.

Lấy mẫu thiên vị. Các cuộc điều tra qua điện thoại bằng cách gọi một mẫu ngẫu nhiên các số trong danh bạ điện thoại được công bố công khai. Đương nhiên việc này loại trừ một cách hệ thống những người có số điện thoại chưa niêm yết, số điện thoại di động và những người không thể trả lời điện thoại (ví dụ, họ đang làm việc) khi cuộc điều tra đang được tiến hành. Điều này cũng bao gồm một số lượng không cân xứng của những người sử dụng điện thoại bàn và những người ở nhà trong nhiều ngày, chẳng hạn như những người thất nghiệp, người tàn tật và người già.

Tương tự như vậy, các cuộc khảo sát trực tuyến có xu hướng nhận được một số lượng không cân xứng sinh viên và những người trẻ tuổi, những người thường xuyên trên

Internet; cũng loại trừ một cách hệ thống những người bị hạn chế hoặc không có điều kiện, khả năng truy cập máy tính hay Internet, chẳng hạn như người nghèo và người cao tuổi. Tương tự như vậy, câu hỏi khảo sát xu hướng loại trừ trẻ em và người mù chữ, không thể đọc, hiểu hoặc không hiểu ý nghĩa câu hỏi.

Một loại khác của thiên vị trong lấy mẫu liên quan đến lấy đối tượng trả lời trong mẫu sai, chẳng hạn như hỏi giáo viên (hoặc cha mẹ) về việc học tập học tập của học sinh mà không hỏi trực tiếp học sinh hoặc hỏi giám đốc điều hành chi tiết về hoạt động trong công ty của họ. Những khiếm khuyết như vậy làm cho mẫu không không đại diện khách thể nghiên cứu và làm hạn chế giá trị các kết luận khái quát được rút ra từ các mẫu thiên vị.

Tính khách quan của thông tin. Nhiều người trả lời có xu hướng tránh những ý kiến tiêu cực hoặc bình luận xấu về bản thân, về người chủ sử dụng lao động, thành viên gia đình hoặc bạn bè. Với câu hỏi tiêu cực, ví như “bạn có nghĩ rằng nhóm tham gia dự án của bạn có vấn đề hay không”, hay “bạn có cho rằng có rất nhiều vấn đề chính trị tại nơi làm việc của bạn” hoặc “bạn có bao giờ bạn tải bất hợp pháp các file nhạc từ Internet”,... các nhà nghiên cứu có thể không nhận được câu trả lời trung thực. Xu hướng này tồn tại trong số những người trả lời thường “bóp méo sự thật” để miêu tả bản thân mình theo cách xã hội mong muốn được gọi là “thiên vị theo mong muốn của xã hội”, gây ảnh hưởng tiêu cực đến giá trị khoa học của các thông tin thu được từ nghiên cứu khảo sát. Trên thực tế, không có cách nào có thể khắc phục thiên vị theo mong muốn của xã hội trong các cuộc khảo sát bằng hỏi, nhưng trong một buổi phỏng vấn, người phỏng vấn sắc sảo có thể để phát hiện ra câu trả lời không phù hợp và đưa ra những câu hỏi thăm dò hoặc sử dụng quan sát cá nhân để bổ sung câu trả lời cần thiết.

Khả năng hồi nhớ. Câu trả lời cho các câu hỏi khảo sát thường phụ thuộc vào động cơ, trí nhớ và khả năng phản ứng của người được hỏi. Đặc biệt, khi hỏi về các sự kiện đã xảy ra trong quá khứ xa xôi, người trả lời có thể không nhớ đầy đủ về động cơ hoặc hành vi của mình hoặc có thể thông tin về sự kiện đó có thể đã mai một theo thời gian và không còn có thể lấy lại được trọn vẹn. Ví dụ, nếu một người được yêu cầu mô tả các sự kiện thời thơ ấu, phản hồi của họ có thể không được chính xác do khó khăn trong việc hồi tưởng. Một cách có thể khắc phục sự khó khăn này là xâu chuỗi các sự kiện cụ thể trong quá khứ, dần dần giúp họ nhớ ra thông tin cần có.

Sai sót trong phương pháp. Sai sót trong phương pháp dùng để chỉ mức độ đồng biến giả mạo giữa các biến độc lập và phụ thuộc được đánh giá tại cùng một điểm (như

trong một nghiên cứu cắt ngang) và sử dụng công cụ tương tự (ví như một bảng câu hỏi). Trong những trường hợp như vậy, hiện tượng đang điều tra có thể không được bóc tách đầy đủ ra khỏi các yếu tố không liên quan. Chúng ta có thể sử dụng các kiểm định thống kê tiêu chuẩn có sẵn để kiểm tra các sai sót trong phương pháp, chẳng hạn như kiểm định nhân tố đơn của Harmon (Podsakoff và các tác giả 2003), kỹ thuật biến thiên thị trường của Lindell và Whitney (2001),... Sai sót này có thể tránh được nếu các biến độc lập và phụ thuộc được đo tại các điểm khác nhau bằng cách sử dụng một thiết kế khảo sát theo chiều dọc; đánh giá các biến bằng các phương pháp khác nhau, chẳng hạn như sử dụng hệ điều hành vi tính để ghi lại sự chuyển đổi của các biến phụ thuộc và các biến độc lập được đánh giá qua bảng câu hỏi.

Chương 10

NGHIÊN CỨU THỰC NGHIỆM

Nghiên cứu thực nghiệm (experimental research)^(*) thường được gọi là “tiêu chuẩn vàng” trong các thiết kế nghiên cứu và là một trong số những thiết kế nghiên cứu khắt khe nhất. Trong thiết kế này, một hay nhiều biến độc lập được điều chỉnh bởi nhà nghiên cứu (bước thiết lập tác động), các khách thể nghiên cứu được bố trí một cách ngẫu nhiên dưới những mức độ tác động khác nhau (bố trí ngẫu nhiên), sau đó kết quả của quá trình tác động tới các biến phụ thuộc sẽ được quan sát, đánh giá. Lợi thế đặc biệt của nghiên cứu thực nghiệm đó là giá trị nhân quả (internal validity), bởi nó có khả năng xác định rõ ràng mối liên hệ giữa nguyên nhân và kết quả thông qua quá trình điều chỉnh sự tác động, cùng với khả năng kiểm soát những ảnh hưởng không mong muốn của các biến liên quan.

Nghiên cứu thực nghiệm phù hợp nhất đối với nghiên cứu giải thích (không phải nghiên cứu mô tả hay nghiên cứu thăm dò). Bởi vì mục đích của nghiên cứu giải thích là để tìm hiểu các mối quan hệ nhân quả. Nó cũng tương đối phù hợp cho những nghiên cứu chứa một số lượng khá nhỏ các biến độc lập được định nghĩa rõ ràng và nhà nghiên cứu có thể điều chỉnh hoặc kiểm soát được các biến này. Nghiên cứu thực nghiệm có thể được thực hiện trong phòng thí nghiệm hay ở ngoài thực địa. Nghiên cứu thí nghiệm (laboratory experiment) được thực hiện trong các phòng thí nghiệm (hay trong môi trường nhân tạo) thường có giá trị nhân quả cao, nhưng đổi lại, giá trị ngoại tại (external validity) lại thấp do các phòng thí nghiệm nhân tạo nơi các nghiên cứu được tiến hành có thể không phản ánh chân thực thực tiễn khách quan.

Nghiên cứu thực nghiệm thực địa (field experiment) được thực hiện trong một không gian thực tế (không chứa các điều kiện nhân tạo), ví dụ như trong một tổ chức. Thực nghiệm thực địa có thể mang đến những giá trị nhân quả và giá trị ngoại tại cao. Tuy nhiên, khó để có thể thực hiện được một thực nghiệm thực địa bởi những khó khăn liên quan đến việc điều chỉnh những yếu tố tác động cũng như kiểm soát các ảnh hưởng không mong

(*) Trong nghiên cứu xã hội, “experimental research” được các tác giả trong nước dịch là “nghiên cứu thực nghiệm”. “Empirical research” cũng được dịch là “nghiên cứu thực nghiệm”, song “empirical” có nghĩa rộng hơn vì nó còn bao hàm cả kinh nghiệm - sự trải nghiệm tự nhiên hoặc là trải nghiệm cố ý trong một thí nghiệm (experiment). Từ điển Wikipedia định nghĩa: “Empirical research is research using empirical evidence. It is a way of gaining knowledge by means of direct and indirect observation or experience” (N.D).

muốn trong điều kiện một không gian thực tiễn.

Nghiên cứu thực nghiệm có thể được phân chia thành hai loại chính: nghiên cứu thực nghiệm đích thực (true experiment) và nghiên cứu giả (hoặc nguy) thực nghiệm (quasi-experiment). Cả hai loại thiết kế này đều đòi hỏi sự điều chỉnh các tác động, nhưng trong khi thực nghiệm đích thực yêu cầu sự phân bố ngẫu nhiên thì nghiên cứu giả thực nghiệm thì không. Đôi khi chúng ta cũng đề cập đến nghiên cứu phi thực nghiệm (non-experimental research). Tuy nó không thật sự là một thiết kế nghiên cứu nhưng nó là một thuật ngữ bao gồm tất cả các mô hình nghiên cứu không sử dụng việc điều chỉnh tác động hay phân bố ngẫu nhiên, ví dụ như nghiên cứu khảo sát, nghiên cứu quan sát và các nghiên cứu tương quan.

Các khái niệm cơ bản

Nhóm thực nghiệm và nhóm đối chứng. Trong nghiên cứu thực nghiệm, một vài đối tượng sẽ chịu một hoặc một số tác động được phân vào *nhóm thực nghiệm* (treatment group), trong khi một số đối tượng khác thì không, được gọi là *nhóm đối chứng* (control group). Thực nghiệm sẽ được xem là thành công nếu những đối tượng trong nhóm thực nghiệm chứng kiến những kết quả tốt hơn so với những đối tượng trong nhóm đối chứng. Chúng ta có thể sử dụng các mức độ tác động khác nhau, trong trường hợp này chúng ta có thể cần phải có nhiều hơn một nhóm thực nghiệm.

Ví dụ, để kiểm nghiệm hiệu quả của một loại thuốc mới dùng để điều trị bệnh mất trí nhớ, nếu một mẫu những bệnh nhân mất trí nhớ được chia thành ba nhóm, trong đó nhóm thứ nhất sẽ nhận liều lượng thuốc cao, nhóm thứ hai nhận liều lượng thấp và nhóm thứ ba sẽ nhận một lượng thuốc vô hại, không tạo ra bất kỳ tác dụng gì, vậy thì hai nhóm đầu là nhóm thực nghiệm và nhóm thứ ba là nhóm đối chứng. Sau khi sử dụng thuốc được một thời gian, nếu như khả năng nhớ của những người thuộc nhóm thực nghiệm được cải thiện đáng kể thì chúng ta có thể nói rằng loại thuốc mới là hiệu quả. Chúng ta cũng có thể so sánh khả năng nhớ của những người thuộc nhóm nhận được liều lượng cao và thấp để đưa ra nhận sử dụng định liều lượng cao hay thấp thì sẽ hiệu quả hơn.

Điều chỉnh tác động (Treatment manipulation). Các tác động là đặc trưng cơ bản để phân biệt nghiên cứu thực nghiệm với các phương pháp nghiên cứu khác. Điều chỉnh tác động này giúp kiểm soát yếu tố “nguyên nhân” trong mối quan hệ nguyên nhân - kết quả. Về cơ bản, giá trị khoa học của nghiên cứu thực nghiệm phụ thuộc vào khả năng và cách thức điều chỉnh các tác động. Điều chỉnh tác động phải được thử nghiệm trước khi thực

hiện chính thức. Tất cả các phương thức đánh giá được thực hiện trước thi sử dụng các tác động được gọi là các đánh giá tiền thực nghiệm (pretest measure), trong khi đó các phương thức đánh giá được thực hiện sau khi sử dụng các tác động được gọi là các đánh giá hậu thực nghiệm (posttest measures).

Lựa chọn và phân bổ ngẫu nhiên (Random selection and assignment). Lựa chọn ngẫu nhiên là một quá trình hình thành một cách ngẫu nhiên một mẫu nghiên cứu từ các khách thể nghiên cứu. Công việc này được thực hiện rất phổ biến trong các nghiên cứu khảo sát nhằm mục đích đảm bảo rằng mỗi một đơn vị trong số khách thể nghiên cứu có cơ hội được lựa chọn vào mẫu. Tuy nhiên, phân bổ ngẫu nhiên trong nghiên cứu thực nghiệm lại là việc phân phối ngẫu nhiên các đối tượng vào nhóm thực nghiệm hay nhóm đối chứng. Đây là một công việc bắt buộc trước khi áp dụng các tác động trong nghiên cứu thực nghiệm đích thực để đảm bảo rằng các nhóm thực nghiệm là tương tự (tương đương) với nhau và với nhóm đối chứng.

Lựa chọn ngẫu nhiên có liên quan tới việc chọn mẫu, vì vậy, nó có mối liên hệ mật thiết hơn tới giá trị ngoại tại (tính khái quát) của các kết quả nghiên cứu. Trong khi đó, phân bổ ngẫu nhiên lại liên quan tới kỹ thuật thiết kế nghiên cứu, vì thế có ảnh hưởng nhiều hơn tới giá trị nhân quả. Chúng ta có thể đạt được cả lựa chọn ngẫu nhiên và phân bổ ngẫu nhiên cùng một lúc trong một nghiên cứu thực nghiệm đích thực được thiết kế phù hợp và chặt chẽ. Trong khi đó, nghiên cứu giả thực nghiệm chỉ đạt được một trong hai yếu tố này.

Các đe dọa đến giá trị nhân quả (Threats to internal validity). Mặc dù thiết kế thực nghiệm được đánh giá là thiết kế khắt khe nhất trong số các phương pháp nghiên cứu dưới góc độ giá trị nhân quả (nhờ khả năng kiểm soát nguyên nhân và điều chỉnh các tác động), tuy nhiên chúng không hoàn toàn “miễn dịch” với các mối đe dọa đối với giá trị nhân quả. Chúng ta sẽ đề cập đến một số nguy cơ này trong ví dụ về một chương trình cải thiện năng lực học toán của các học sinh phổ thông.

- *Yếu tố lịch sử* là khả năng những kết quả được ghi nhận (từ các biến phụ thuộc) gây ra bởi các yếu tố hay sự kiện lịch sử không liên quan, không phải bởi các tác động được áp dụng trong thực nghiệm. Ví dụ, sự cải thiện điểm số toán học của học sinh sau khi thực áp dụng các tác động xuất phát từ sự chuẩn bị tích cực trước đó của học sinh cho kỳ thi toán cuối năm chứ không phải xuất phát từ việc áp dụng các tác động trong chương trình này.

- *Yếu tố phát triển tự nhiên* (maturation) nói đến khả năng kết quả quan sát được gây

ra bởi sự phát triển tự nhiên của khách thể nghiên cứu. Ví dụ học sinh cải thiện điểm số toán học thực tế là do sự cải thiện năng lực tri thức thực sự của học sinh, chứ không phải là do tác động của chương trình nghiên cứu mang lại.

- *Yếu tố kiểm định* là mối đe dọa xuất hiện khi những phản ứng trong quá trình thực nghiệm chịu ảnh hưởng của các phản ứng tiền kiểm định trước đó. Ví dụ, nếu như học sinh nhớ được các đáp án trong các đánh giá tiền kiểm định, họ có thể nhắc lại chúng trong kỳ kiểm tra sau. Không thực hiện bước tiền kiểm định có thể giúp tránh được mối đe dọa này.

- *Yếu tố đo lường* đề cập đến sự khác biệt về mức độ đo lường trước và sau khi thực nghiệm. Ví dụ các bài kiểm tra đánh giá môn toán trước và sau khi chương trình áp dụng có độ khó khác nhau.

- *Yếu tố mất mát* liên qua đến khả năng vì một lý do nào đó các đối tượng có thể rời bỏ khỏi nghiên cứu làm mất cân đối trong tỉ lệ giữa các nhóm thực nghiệm và nhóm đối chứng. Ví dụ một số học viên bị điểm kém tại kỳ kiểm tra thử có thể bỏ khỏi nghiên cứu. Trong trường hợp này, kết quả của kỳ thi cuối khóa sẽ bị thời phồng bởi sự lấn át của các học sinh có điểm số cao.

- *Sai sót thống kê*: Quá trình thử nghiệm, đánh giá các quan hệ nhân quả thường qua nhiều bước, sử dụng nhiều kỹ thuật thống kê khác nhau. Nếu việc nhập số liệu vào các kỹ thuật này không được thực hiện nhất quán hoặc kỹ thuật được sử dụng không tương thích thì kết quả thu được sẽ không phản ánh bản chất của các mối quan hệ đang nghiên cứu.

Thiết kế thực nghiệm hai nhóm (Two-Group Experimental Designs).

Thiết kế thực nghiệm đích thực đơn giản nhất là thiết kế hai nhóm bao gồm một nhóm thực nghiệm và một nhóm đối chứng. Phương pháp này thật sự lý tưởng cho việc kiểm định ảnh hưởng của một biến độc lập mà biến này có thể điều chỉnh được. Có hai loại thiết kế thực nghiệm hai nhóm chính đó là (1) thiết kế nhóm đối chứng tiền - hậu kiểm định và (2) thiết kế nhóm đối chứng hậu kiểm định. Những thiết kế này được mô tả bằng cách sử dụng một ký hiệu thiết kế tiêu chuẩn hóa. Trong đó, R đại diện cho phân bổ ngẫu nhiên của các đối tượng thành các nhóm, X đại diện cho các tác động được áp dụng đối với nhóm thực nghiệm và O đại diện cho quan sát trước hoặc sau kiểm định đối với biến phụ thuộc (với những ký hiệu khác nhau để phân biệt các quan sát trước và sau kiểm định đối với các nhóm thực nghiệm và đối chứng).

Thiết kế nhóm đối chứng tiền - hậu kiểm định (Pretest-posttest control group design).

Trong thiết kế này, người tham gia vào nghiên cứu được phân ngẫu nhiên vào nhóm đối chứng và nhóm thực nghiệm. Ngay sau đó, giá trị các biến độc lập ở cả hai nhóm sẽ được đo lường (bước tiền kiểm định). Tiếp theo, nhóm thực nghiệm được tương tác với thử nghiệm. Sau quá trình tương tác, các biến phụ thuộc lại được đo lường (bước hậu kiểm định). Các ký hiệu của thiết kế này được thể hiện trong Hình 10.1 .

R	O ₁	X	O ₂	(Nhóm thực nghiệm)
R	O ₃		O ₄	(Nhóm đối chứng)

Hình 10.1. Thiết kế nhóm đối chứng tiền - hậu kiểm định

Hiệu quả tác động E sau quá trình áp dụng các thử nghiệm trong thiết kế tiền - hậu kiểm định này được tính bằng hiệu số giữa kết quả kiểm định trước và sau của nhóm thực nghiệm và nhóm đối chứng:

$$E = (O_2 - O_1) - (O_4 - O_3)$$

Phân tích thống kê của thiết kế này bao gồm phân tích phương sai đơn giản (analysis of variance - ANOVA) giữa nhóm thực nghiệm và nhóm đối chứng. Trong thiết kế tiền - hậu kiểm định, có một số yếu tố có thể làm giảm giá trị nhân quả. Chẳng hạn như diễn biến tự nhiên của đối tượng nghiên cứu, ảnh hưởng của bước tiền thử nghiệm và sai sót trong kỹ thuật thống kê. Điều này là do các yếu tố này ngẫu nhiên gây ra những tác động tương tự đến cả hai nhóm thực nghiệm và đối chứng. Những đe dọa này có thể được khắc phục thông qua sự phân bố ngẫu nhiên. Tuy nhiên, các mối đe dọa tới giá trị nhân quả khác vẫn có khả năng tồn tại. Ví dụ, yếu tố mất mát (như đề cập ở trên) có thể xuất hiện nếu tỷ lệ người rút khỏi hai nhóm khác nhau, hoặc có sự khác biệt trong phương pháp đo lường trong tiền kiểm định và hậu kiểm định (đặc biệt là khi bước tiền kiểm định đưa ra những chủ đề hoặc nội dung kiểm tra khác lạ).

Thiết kế nhóm đối chứng hậu kiểm định (Posttest-only control group design).

Thiết kế này là phiên bản giản lược của thiết kế tiền - hậu kiểm định, trong đó, đo lường tiền kiểm định bị lược bỏ. Mô hình thiết kế này được trình bày ở Hình 10.2.

R	X	O ₁	(Nhóm thực nghiệm)
R		O ₂	(Nhóm đối chứng)

Hình 10.2. Thiết kế nhóm đối chứng hậu kiểm định

Hệ quả tác động (E) được đánh giá đơn giản chỉ là sự khác biệt trong kết quả hậu kiểm định giữa hai nhóm:

$$E = (O_1 - O_2)$$

Kỹ thuật phân tích thống kê phù hợp với thiết kế hậu kiểm định này cũng là kỹ thuật phân tích phương sai hai nhóm (two-group analysis of variance - ANOVA). Sự đơn giản của thiết kế hậu kiểm định làm cho nó có khả năng đạt được giá trị nhân - quả cao hơn so với thiết kế tiền - hậu kiểm định. Thiết kế này mang đến khả năng kiểm soát được các mối đe dọa đến từ diễn biến tự nhiên của đối tượng nghiên cứu, ảnh hưởng của bước tiền thử nghiệm, sai sót trong kỹ thuật thống kê và tương tác tiền - hậu kiểm định. Tuy vậy, mối đe dọa mất mát vẫn tiếp tục tồn tại.

Thiết kế phương sai (Covariance designs). Đôi khi, việc đo lường các biến phụ thuộc bị ảnh hưởng bởi các biến ngoại sinh được gọi là đồng biến (hay biến nhiễu - Covariates). Đồng biến là những biến không phải là trọng tâm của nghiên cứu nhưng lại có những hiệu ứng nhất định lên mối quan hệ nhân quả của biến độc lập và biến phụ thuộc. Vì vậy hiệu ứng này cần phải được nhận diện và kiểm soát chặt chẽ. Các thiết kế thử nghiệm trước không đề cập đến vấn đề này. Thiết kế phương sai (cũng được gọi là thiết kế đồng biến) là một phiên bản đặc biệt của thiết kế nhóm đối chứng tiền - hậu kiểm định với bước tiền kiểm chỉ đo các đồng biến, mà không phải là các biến phụ thuộc. Các ký hiệu của thiết kế này được thể hiện trong Hình 10.3, trong đó C đại diện cho các đồng biến.

R	C	X	O ₁	(Nhóm thực nghiệm)
R	C		O ₂	(Nhóm đối chứng)

Hình 10.3. Thiết kế phương sai

Bởi vì công cụ tiền kiểm không đo biến phụ thuộc mà đo các đồng biến nên hiệu quả tác động E được tính bằng hiệu số giữa nhóm thử nghiệm và nhóm đối chứng trong quá trình hậu kiểm.

$$E = (O_1 - O_2)$$

Do có sự xuất hiện của các đồng biến, kỹ thuật thống kê phù hợp là phân tích hiệp

phương sai hai nhóm (two-group analysis of covariance). Thiết kế này kế thừa tất cả ưu điểm của thiết kế nhóm đối chứng hậu kiểm định. Tuy vậy, nó cũng có thể phát triển thành một thiết kế nhóm đối chứng tiền - hậu kiểm định.

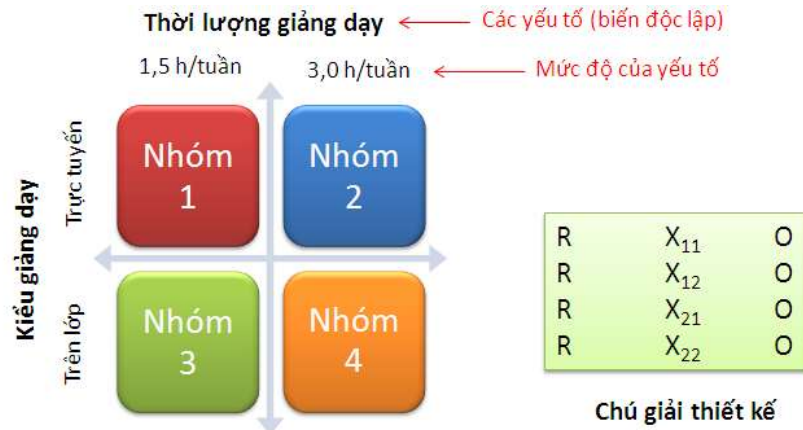
Thiết kế giai thừa (Factorial designs)

Thiết kế hai nhóm không phù hợp đối với những nghiên cứu đòi hỏi sự điều chỉnh hai hay nhiều hơn hai biến độc lập (vì một biến cần 2 nhóm thực nghiệm và đối chứng). Trong trường hợp này, chúng ta cần có thiết kế bốn nhóm hoặc nhiều hơn. Thiết kế này khá phổ biến trong nghiên cứu thực nghiệm, được gọi là thiết kế giai thừa (factorial designs). Mỗi biến độc lập trong thiết kế này được gọi là một giai thừa (factor) và mỗi phần nhỏ của giai thừa được gọi là một mức (level). Thiết kế giai thừa cho phép nhà nghiên cứu kiểm tra không chỉ tác động riêng lẻ của mỗi biến độc lập lên biến phụ thuộc tương ứng (gọi là tác động chính) mà còn tác động chung của các biến độc lập lên từng biến phụ thuộc riêng lẻ (được gọi là tác động tương tác).

Thiết kế giai thừa cơ bản nhất là thiết kế 2×2 , gồm hai nhóm thử nghiệm, mỗi nhóm thử nghiệm có hai mức (chẳng hạn như thấp/cao hoặc có mặt/vắng mặt). Ví dụ, bạn muốn so sánh kết quả học tập từ hai phương pháp giảng dạy khác nhau (trên lớp và trực tuyến). Bạn muốn kiểm tra xem kết quả sẽ thay đổi như thế nào khi áp dụng thời lượng giảng dạy khác nhau (1,5 hoặc 3 giờ mỗi tuần). Trong trường hợp này bạn có hai thừa số: phương pháp giảng dạy và thời lượng giảng dạy (thể hiện trong Hình 10.4). Nếu bạn muốn bổ sung thêm mức thứ ba cho thời lượng giảng dạy (ví dụ 6 giờ/tuần), thì giai thừa này (thời lượng giảng dạy) sẽ có ba mức và bạn sẽ có một thiết kế giai thừa 2×3 . Trong khi đó, nếu bạn muốn bổ sung thêm một giai thừa thứ ba chẳng hạn như giai thừa phương pháp học tập theo nhóm (với hai mức áp dụng và không áp dụng), như vậy thì bạn sẽ có một thiết kế giai thừa $2 \times 2 \times 2$. Trong ký hiệu này thì mỗi một con số đại diện cho một giai thừa (ví dụ 2×3 nghĩa là có 2 giai thừa, $2 \times 2 \times 2$ tức là có 3 giai thừa), trong khi đó giá trị của con số đó chỉ ra giai thừa đó có bao nhiêu mức.

Các thiết kế giai thừa cũng có thể được mô tả thông qua ký hiệu thiết kế như được trình bày trong hộp phía bên phải của Hình 10.4. Ký hiệu R nghĩa là phân bố ngẫu nhiên, X đại diện cho các nhóm thử nghiệm (các con số bên dưới X cho biết các mức của giai thừa), O đại diện cho kết quả quan sát các biến phụ thuộc. Chú ý rằng, thiết kế 2×2 sẽ có bốn nhóm thử nghiệm, tương đương với tổng số bốn lần kết hợp của hai mức cho mỗi giai thừa. Như vậy thiết kế 2×3 sẽ có sáu nhóm thử nghiệm và thiết kế $2 \times 2 \times 2$ sẽ có tám nhóm.

Như là một nguyên tắc phổ quát, mỗi ngăn trong một thiết kế giai thừa nên có số lượng ít nhất là 20 khách thể nghiên cứu trong mẫu (nguyên tắc này xuất phát từ tính toán của Cohen về kích thước mẫu trung bình). Do vậy, trong thiết kế giai thừa $2 \times 2 \times 2$, số lượng khách thể nghiên cứu tối thiểu trong mẫu phải là 160, với 20 khách thể nằm trong một ngăn. Vì vậy, chi phí cho việc thu thập dữ liệu sẽ tăng lên đáng kể nếu tăng số lượng giai thừa hoặc tăng số mức cho mỗi giai thừa.



Hình 10.4. Thiết kế giai thừa 2×2

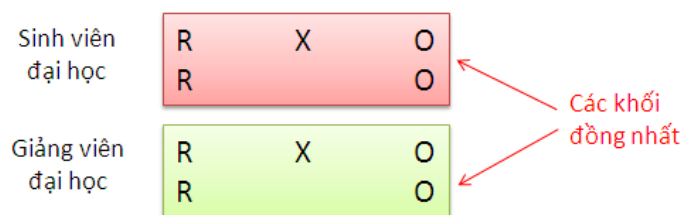
Trên thực tiễn đôi khi do thiếu nguồn lực về thời gian, nhân lực, kinh phí mà một vài ngăn không được tiếp nhận bất kỳ một công cụ thử nghiệm nào. Tình huống này được gọi là thiết kế giai thừa không hoàn thiện và những thiết kế này sẽ làm giảm giá trị của các kết luận rút ra từ các giai thừa không hoàn chỉnh. Trong một thiết kế giai thừa, tác động chính (main effect) được cho là tồn tại nếu biến phụ thuộc chỉ ra một sự khác biệt rõ ràng khi chúng ta điều chỉnh các mức khác nhau của giai thừa (biến độc lập). Khi biến phụ thuộc không cho thấy sự khác biệt nào giữa các mức của giai thừa, ta gọi đó là trường hợp vô hiệu. Trong ví dụ trên, bạn có thể tìm thấy tác động chính của phương pháp giảng dạy, thời lượng giảng dạy hoặc của cả hai đối với kết quả học tập.

Một tác động tương tác (interaction effect) được coi là tồn tại khi tác động chính của một giai thừa thay đổi theo các mức khác nhau của giai thừa khác. Trong ví dụ trên tác động của phương pháp giảng dạy tới kết quả học tập mạnh mẽ hơn khi áp dụng thời lượng giảng dạy là ba giờ/tuần so với khi áp dụng thời lượng 1,5 giờ/tuần. Vậy thì chúng ta có thể nói rằng có một tác động tương tác giữa phương pháp giảng dạy và thời lượng giảng dạy đối với kết quả học tập. Chú ý rằng có trường hợp sự xuất hiện của các tác động tương tác có thể sẽ lấn át và làm cho tác động chính không còn ý nghĩa.

Thiết kế thực nghiệm hỗn hợp (Hybrid experimental designs).

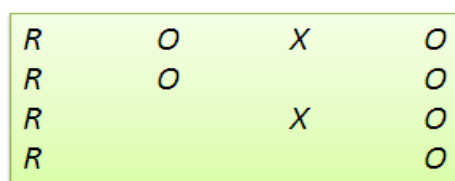
Thiết kế hỗn hợp được hình thành bởi kết hợp các thiết kế chính hiện có. Ba thiết kế hỗn hợp chính là: thiết kế khối ngẫu nhiên, thiết kế nhóm tứ Solomon và thiết kế sao chép hoán đổi.

Thiết kế khối ngẫu nhiên (Randomized block design). Đây là một biến thể của thiết kế nhóm đối chứng tiền - hậu kiểm định hoặc thiết kế nhóm đối chứng hậu kiểm định, trong đó các khách thể nghiên cứu được phân vào các nhóm tương đối đồng nhất (gọi là các khối) từ đó thử nghiệm được thực hiện lại. Ví dụ, nếu bạn muốn kiểm tra lại kết quả từ thiết kế hậu kiểm định trong số sinh viên đại học và các giảng viên (hai khối đồng nhất), thì những người được nghiên cứu ở cả hai khối được ngẫu nhiên phân vào nhóm thử nghiệm (sẽ tiếp nhận cùng một công cụ thử nghiệm) và nhóm đối chứng (Hình 10.5). Mục đích của thiết kế này là để giảm tác động của những yếu tố gây nhiễu, nhờ đó bản chất mối quan hệ nhân quả có thể được phát hiện chính xác hơn.



Hình 10.5. Thiết kế khối ngẫu nhiên

Thiết kế nhóm tứ Solomon (Solomon four - group design). Trong thiết kế này, thành viên trong mẫu được phân vào hai nhóm thử nghiệm và hai nhóm đối chứng. Một nhóm thử nghiệm và một nhóm đối chứng sẽ được đo lường tiền kiểm, hai nhóm còn lại thì không. Thiết kế này là một sự kết hợp giữa thiết kế tiền - hậu kiểm định và hậu kiểm định, nhằm mục đích phát hiện những tác động không mong muốn của quá trình tiền kiểm tới quá trình hậu kiểm. Những tác động này có thể xảy ra trong thiết kế tiền - hậu kiểm định mà không xảy ra trong hậu kiểm định. Bằng việc kết hợp cả hai loại thiết kế này, thiết kế nhóm tứ Solomon có thể phát hiện ra những tác động không mong muốn đó. Ký hiệu của thiết kế này được thể hiện trong Hình 10.6.



Hình 10.6. Thiết kế nhóm tứ Solomon

Thiết kế sao chép hoán đổi (Switched replication design). Đây là thiết kế hai nhóm thực hiện trong hai giai đoạn với ba lần đo lường. Nhóm thử nghiệm trong giai đoạn thứ nhất sẽ đóng vai trò là nhóm đối chứng trong giai đoạn thứ hai và nhóm đối chứng trong giai đoạn thứ nhất trở thành nhóm thử nghiệm trong giai đoạn hai, như được mô tả trong hình 10.7. Nói cách khác, thiết kế ban đầu được lặp lại với vai trò của các nhóm được hoán đổi. Kết thúc nghiên cứu, tất cả khách thể tham gia nghiên cứu đều được tiếp nhận công cụ thử nghiệm có thể là trong giai đoạn 1 hoặc giai đoạn 2. Thiết kế này sẽ khả thi nhất khi thực hiện trong các tổ chức nơi mà các chương trình công tác được thực hiện theo từng giai đoạn và lặp lại trong những khoảng thời gian nhất định.

R	O	X	O	O	O
R	O	O	X	O	O

Hình 10.7. Thiết kế sao chép hoán đổi

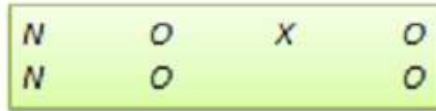
Thiết kế giả thực nghiệm (Quasi-experimental designs).

Thiết kế giả thực nghiệm (hay nguy thực nghiệm) có nhiều nét tương đồng với các thiết kế thực nghiệm đích thực nhưng thiếu đi một yếu tố chính là phân bổ ngẫu nhiên. Việc thiếu đi sự phân bổ ngẫu nhiên tạo ra khả năng có những nhóm không tương xứng. Ví dụ, khi chia các lớp học với toàn bộ học sinh của mỗi lớp trong một trường thành hai nhóm: thực nghiệm và đối chứng. Việc phân nhóm này có thể xuất hiện tình huống một nhóm sẽ có khả năng học tập tốt hơn nếu như nhóm này có các lớp mà học kỳ trước được giảng dạy bởi những giáo viên tốt hơn. Điều này sẽ mang đến khả năng thiên vị trong lựa chọn. Chính vì vậy, giả thực nghiệm không được đánh giá cao về giá trị nhân quả so với thực nghiệm đích thực.

Hàng loạt yếu tố khác có thể đe dọa giá trị nội (nhân quả) của thiết kế nguy thử nghiệm: diễn tiến tự nhiên của khách thể tham gia nghiên cứu, yếu tố lịch sử (trước khi thử nghiệm người tham gia nghiên cứu đã chịu một số tác động mà nhà nghiên cứu không lường trước được), công cụ thử nghiệm (một số người phản ứng với công cụ thử nghiệm theo cách không thông thường), sai số trong sử dụng các kỹ thuật thống kê, yếu tố mất mát (tỉ lệ người tham gia rút khỏi nghiên cứu khác nhau trong nhóm thử nghiệm và nhóm đối chứng). Chính vì những mối đe dọa này, nhìn chung thiết kế giả thực nghiệm không được khuyến khích áp dụng.

Nhiều thiết kế thử nghiệm đích thực có thể bị chuyển đổi thành giả thực nghiệm. Ví

dụ như thiết kế nhóm đối chứng tiền - hậu kiểm định có thể chuyển thành thiết kế nhóm phi tương đương (nonequivalent groups design - NEGD) như Hình 10.8, khi phân bố ngẫu nhiên R được thay thế bởi phân bố phi tương đương (không ngẫu nhiên) N. Tương tự như vậy, thiết kế sao chép hoán đổi có thể trở thành thiết kế sao chép hoán đổi phi tương đương (non-equivalent switched replication design) như mô tả trong Hình 10.9.



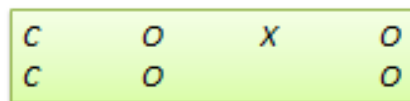
Hình 10.8. Thiết kế nhóm không tương đương



Hình 10.9. Thiết kế sao chép hoán đổi không tương đương

Ngoài ra, có một số thiết kế phi tương đương đặc biệt khi chúng không phát triển trên trên mô hình của các thiết kế thử nghiệm đích thực. Một vài thiết kế này hữu ích trong khoa học xã hội và sẽ được trình bày sau đây.

Thiết kế hồi quy gián đoạn (Regression - discontinuity (RD) design). Đây là thiết kế tiền hậu kiểm định phi tương đương, trong đó người tham gia nghiên cứu được phân bổ vào nhóm thử nghiệm và nhóm đối chứng dựa trên một ngưỡng được xác định từ trước. Ví dụ, những bệnh nhân với thể trạng rất yếu được phân vào nhóm thử nghiệm để kiểm tra tác dụng của một loại thuốc mới hoặc một pháp đồ điều trị mới. Trong khi đó những bệnh nhân có thể trạng trung bình được phân vào nhóm đối chứng. Một ví dụ khác là những học sinh có học lực kém được lựa chọn vào một chương trình giảng dạy cải tiến nhằm mục đích cải thiện học lực của họ, trong khi đó những học viên có học lực tốt không được lựa chọn vào chương trình này. Ký hiệu của thiết kế này được trình bày dưới đây, với C đại diện cho ngưỡng phân bổ.

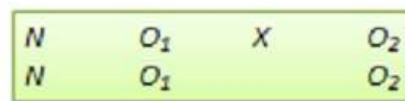


Hình 10.10. Thiết kế hồi quy gián đoạn

Vì lí do áp đặt ngưỡng phân bổ, kết quả thử nghiệm có thể bị ảnh hưởng bởi đặc trưng của ngưỡng, điều này mang đến một mối đe dọa mới với giá trị nhân quả. Tuy vậy, việc áp dụng ngưỡng sẽ là hữu ích khi mà nguồn lực cho đề tài có hạn, và những người

nằm trong ngưỡng này là những người có nhu cầu cấp thiết hơn được tiếp nhận kết quả nghiên cứu khi nghiên cứu thành công. Điểm số của nhóm đối chứng trong thiết kế RD không được xem là tiêu chuẩn để so sánh với điểm số của nhóm thử nghiệm bởi vì có sự phân bố không tương đương giữa hai nhóm. Thay vào đó, nếu không có sự khác biệt (gián đoạn) giữa điểm số của nhóm đối chứng trước và sau thử nghiệm, trong khi đó có sự khác biệt rõ ràng trong nhóm thử nghiệm, thì sự khác biệt này được xem như bằng chứng cho tác dụng của công cụ thử nghiệm.

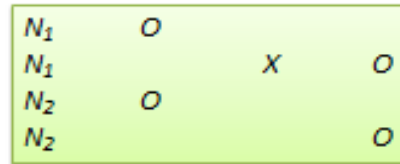
Thiết kế tiền kiểm đại diện (Proxy pretest design). Thiết kế này, như nhìn thấy trong Hình 10.11, trông rất giống thiết kế tiền hậu kiểm định trước đó. Tuy vậy nó có một điểm khác biệt chính đó là điểm số tiền kiểm lại được thu nhận sau thời điểm thử nghiệm được thực hiện. Một ứng dụng phổ biến của thiết kế này là khi nhà nghiên cứu muốn đánh giá hiệu quả của một chương trình (ví dụ một chương trình giáo dục) sau khi chương trình này đã triển khai và do vậy lúc này không thể thực hiện tiền kiểm được. Trong tình huống này, giải pháp tốt nhất thường là sử dụng kết quả đã được đo lường trước đó (ví dụ như điểm số của học sinh trước khi chương trình được triển khai) như là đại diện cho điểm số tiền kiểm.



Hình 10.11. Thiết kế tiền kiểm đại diện

Thiết kế tiền - hậu kiểm riêng (Separate pretest-posttest samples design). Thiết kế này sẽ hữu ích nếu như vì lí do nào đó mà việc thu thập thông tin từ những người tham gia nghiên cứu không thực hiện được cả trước và sau thử nghiệm. Ví dụ, bạn muốn đánh giá mức độ hài lòng của khách hàng về một dịch vụ trực tuyến mới, mà dịch vụ này chỉ được triển khai trong ở thành phố A mà không ở thành phố B. Trong trường hợp này khách hàng ở thành phố A đóng vai trò là nhóm thử nghiệm, còn khách hàng ở thành phố B trở thành nhóm đối chứng. Tại thành phố A, nếu bạn không thể thực hiện các đánh giá tiền kiểm và hậu kiểm cho cùng một nhóm khách hàng, bạn có thể đo mức độ hài lòng của nhóm khách hàng nào đó, sau đó triển khai thực hiện dịch vụ mới này, và đo độ hài lòng (với nhóm khách hàng khác) sau khi dịch vụ kết thúc. Việc đo mức độ hài lòng của khách hàng cũng được thực hiện với nhóm đối chứng ở thành phố B (dịch vụ trực tuyến không triển khai) tại cùng thời điểm thực hiện tại thành phố A. Thiết kế này không thật sự mạnh bởi vì bạn không thể kiểm tra được sự thay đổi trong mức độ hài lòng của cùng nhóm khách hàng

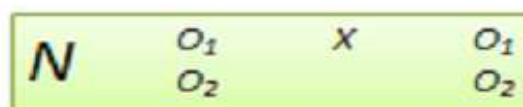
trước và sau khi thực hiện dịch vụ, vì vậy bạn chỉ có thể đánh giá sự thay đổi này một cách tương đối mà thôi. Tuy có giá trị nhân quả thấp hơn, nó vẫn hữu ích trong việc thu thập những dữ liệu nguy thử nghiệm khi mà chúng ta không thể thu được dữ liệu tiền kiểm và hậu kiểm cho cùng một mẫu khách thể nghiên cứu.



Hình 10.12. Thiết kế mẫu tiền - hậu kiểm riêng

Thiết kế biến độc lập phi tương đương (Nonequivalent dependent variable (NEDV) design). Đây là một thiết kế nguy thử nghiệm tiền kiểm một nhóm, với hai biến phụ thuộc. Biến thứ nhất, về mặt lý thuyết, sẽ bị ảnh hưởng bởi công cụ thử nghiệm và biến thứ hai thì không. Ví dụ, bạn muốn đánh giá tính hiệu quả của một chương trình giảng dạy môn giải tích mới cho học sinh trung học phổ thông. Chương trình này là một công cụ thử nghiệm và được cho là sẽ mang đến những tác động tích cực trong kết quả học tập giải tích vào cuối năm. Tuy nhiên, kết quả học tập môn đại số vào cuối năm vẫn có thể cải thiện bởi vì những yếu tố ngoại lai chẳng hạn như yếu tố lịch sử hoặc yếu tố diễn tiến tự nhiên. Như vậy thì kết quả học tập môn đại số cuối năm có thể được sử dụng như điểm đối chứng, trong khi đó điểm số đầu năm môn giải tích được xem như điểm thử nghiệm. Ký hiệu cho thiết kế này, như được thấy trong Hình 10.13, chỉ ra rằng một nhóm duy nhất N được đo tiền kiểm O1 và hậu kiểm O2 cho hai môn giải tích và đại số cho cùng một nhóm học sinh. Thiết kế này có giá trị nhân quả thấp, nhưng nó chỉ cần sử dụng một nhóm duy nhất vừa làm nhóm thử nghiệm vừa làm nhóm đối chứng.

Một biến thể thú vị của thiết kế NEDV này là một thiết kế NEDV cộng hưởng. Nó sử dụng nhiều biến phụ thuộc khác nhau và dùng một lý thuyết để giải thích mức độ ảnh hưởng bởi công cụ thử nghiệm đối với mỗi biến riêng lẻ. Và như vậy nhà nghiên cứu có thể đánh giá tổng hợp mức độ ảnh hưởng của tất cả các biến. Phương pháp cộng hưởng này là một công cụ hữu hiệu để cải thiện giá trị nhân quả trong thiết kế NEDV đơn thuần.



Hình 10.13. Thiết kế biến độc lập phi tương đương (NEDV)

Những lưu ý trong nghiên cứu thực nghiệm

Nghiên cứu thực nghiệm là một trong số những thiết kế nghiên cứu khó nhất và nên được thực hiện với sự tập trung cao độ. Loại nghiên cứu này đối mặt với vô số vấn đề về phương pháp luận. Đầu tiên, mặc dù nghiên cứu thực nghiệm đòi hỏi phải sử dụng lý thuyết để hình thành nên các giả thuyết phục vụ quá trình kiểm định, nhưng gần đây một số lượng đáng kể các nghiên cứu thực nghiệm là phi lý thuyết. Không có lý thuyết, các giả thuyết có xu hướng ngẫu hứng, có thể thiếu logic và không có nghĩa.

Thứ hai, nhiều công cụ đánh giá được sử dụng trong nghiên cứu thực nghiệm không được kiểm định về độ tin cậy và độ giá trị và không thể so sánh được trong các nghiên cứu khác nhau. Thứ ba, nhiều nghiên cứu sử dụng các thiết kế nghiên cứu không phù hợp, ví dụ như thiết kế các biến phụ thuộc không liên quan, không có kết quả tương tác, không có sự kiểm soát các tác động và tác động không tương thích đối với các nhóm thực nghiệm. Kết quả từ những nghiên cứu như vậy thường có giá trị nhân quả không cao và mang đến nhiều hoài nghi. Ví dụ, các sinh viên đại học được đề nghị đóng vai các nhà quản lý thị trường và được đề nghị thực hiện nhiệm vụ phân bổ ngân sách phức tạp, trong khi đó họ không có kinh nghiệm hay chuyên môn sâu về công việc này. Việc sử dụng một nhiệm vụ không phù hợp này sẽ làm giảm giá trị nhân quả (ví dụ sự thể hiện của sinh viên là không đúng thực tế), mang đến những phát hiện không giải thích được và không có ý nghĩa, làm cho việc phối kết hợp các phát hiện trong các nghiên cứu trở nên bất khả thi.

Thiết kế các tác động (công cụ thử nghiệm) đúng đắn là một nhiệm vụ vô cùng quan trọng trong nghiên cứu thực nghiệm bởi vì các công cụ thử nghiệm là linh hồn và “lẽ sống” của phương pháp thực nghiệm, do đó không bao giờ được phép lơ là hay vội vàng trong nhiệm vụ này. Để có được một tác động phù hợp, nhà nghiên cứu phải thực hiện các bước củng cố và thí điểm, kiểm tra tính tương thích của các công cụ, nếu có nghi ngờ, áp dụng các công cụ đơn giản hơn, quen thuộc hơn đối với người trả lời.

Nói tóm lại, chương này đã giới thiệu các khái niệm cơ bản trong phương pháp nghiên cứu thực nghiệm cũng như hàng loạt các thiết kế nghiên cứu thực nghiệm đích thực và giả thực nghiệm khác nhau. Mặc dù các thiết kế này mang đến những sự khác biệt lớn trong giá trị nhân quả, nhưng không nên bỏ qua các thiết kế có giá trị nhân quả thấp bởi vì hoặc đôi khi chúng hữu ích trong một số tình huống cụ thể.

Chương 11

NGHIÊN CỨU TRƯỜNG HỢP

Nghiên cứu trường hợp (Case study) là phương pháp nghiên cứu chuyên sâu về một hiện tượng trong một khoảng thời gian liên tục và trong bối cảnh tự nhiên của hiện tượng đó. Nhiều phương pháp thu thập dữ liệu, như phỏng vấn, quan sát và nghiên cứu dữ liệu thứ cấp, có thể được sử dụng đồng thời và những suy luận về hiện tượng quan tâm thường phong phú, chi tiết và phù hợp với ngữ cảnh. Nghiên cứu trường hợp có thể được sử dụng theo hướng tiếp cận thực chứng cho mục đích kiểm nghiệm lý thuyết hoặc tiếp cận diễn giải để xây dựng lý thuyết. Phương pháp này phổ biến trong nghiên cứu kinh doanh hơn trong các ngành khoa học xã hội khác.

Nghiên cứu trường hợp có những điểm mạnh đặc trưng hấp dẫn hơn các phương pháp nghiên cứu khác như thực nghiệm và nghiên cứu khảo sát. *Thứ nhất*, nghiên cứu trường hợp có thể được sử dụng cho cả xây dựng lý thuyết và cả kiểm nghiệm lý thuyết, trong khi đó các phương pháp thực chứng khác chỉ có thể được sử dụng để kiểm nghiệm lý thuyết. Bên cạnh đó, trong nghiên cứu trường hợp diễn giải, không cần phải hình thành các phạm trù từ trước, mà chúng có thể xuất hiện từ các dữ liệu thu được trong quá trình nghiên cứu. *Thứ hai*, các câu hỏi nghiên cứu có thể được sửa đổi trong quá trình nghiên cứu khi nhận thấy các câu hỏi ban đầu ít liên quan hoặc không nổi bật. Điều này gần như là không thể thực hiện được trong các phương pháp thực chứng sau khi dữ liệu được thu thập. *Thứ ba*, nghiên cứu trường hợp có thể giúp tạo ra các giải thích phong phú hơn, nhiều màu sắc hơn và xác thực hơn về hiện tượng quan tâm so với hầu hết các phương pháp nghiên cứu khác nhờ khả năng thu hoạch hàng loạt dữ liệu phong phú theo ngữ cảnh. *Thứ tư*, hiện tượng quan tâm có thể được nghiên cứu từ quan điểm của nhiều người tham gia và sử dụng nhiều cấp độ phân tích (ví dụ cả cấp độ cá nhân và tổ chức).

Bên cạnh đó, nghiên cứu trường hợp cũng có một số điểm yếu cố hữu. Do nó không có sự đối chứng thực nghiệm, giá trị nhân quả của các suy luận thường thấp. Đây cũng là một điểm yếu chung cho tất cả các phương pháp nghiên cứu, ngoại trừ phương pháp thực nghiệm. Tuy nhiên, như các mô tả sau này, vấn đề kiểm định này có thể được giải quyết trong nghiên cứu trường hợp bằng cách sử dụng “đối chứng tự nhiên”. *Thứ hai*, chất lượng của các kết luận thu được từ nghiên cứu trường hợp phụ thuộc rất nhiều vào khả năng tổng

hợp, phân tích của nhà nghiên cứu. Một nhà nghiên cứu có kinh nghiệm có thể nhận ra các khái niệm và từ đó xây dựng mô hình lý thuyết trong cùng một dữ liệu. Nhưng một nhà nghiên cứu mới có thể bỏ lỡ. Do đó, những phát hiện có được đôi khi bị chỉ trích là chủ quan. *Cuối cùng*, bởi những kết luận từ nghiên cứu trường hợp là từ những bối cảnh rất cụ thể, có thể là khó khăn để khái quát tới các bối cảnh khác.

Chúng ta cũng cần nhận ra sự khác nhau giữa *nghiên cứu trường hợp* với *mô tả trường hợp* như các tình huống được thảo luận trong các lớp học kinh doanh tại trường Đại học Harvard. Mô tả trường hợp thường mô tả chi tiết một vấn đề của tổ chức với mục đích khuyến khích thảo luận trong lớp học và tư duy phê phán giữa các học sinh hoặc phân tích cách thức một tổ chức xử lý một vấn đề cụ thể. Trong khi đó, nghiên cứu trường hợp là một kỹ thuật nghiên cứu chính thức, sử dụng các phương pháp khoa học để đưa ra những giải thích cho các hiện tượng của tổ chức.

Nghiên cứu trường hợp là một phương pháp khó khi nó đòi hỏi nhà nghiên cứu phải có kỹ năng nghiên cứu phức tạp, do đó nó có thể dễ dàng được thực hiện không đúng. Benbasat và các tác giả (1987)⁸ mô tả năm vấn đề thường gặp trong các nghiên cứu trường hợp. *Thứ nhất*, nhiều nghiên cứu trường hợp bắt đầu mà không có câu hỏi nghiên cứu cụ thể; do đó kết thúc mà không đưa ra bất kỳ câu trả lời cụ thể hoặc suy luận sâu sắc nào. *Thứ hai*, hướng tiếp cận vào trường hợp cần nghiên cứu thường được dựa trên khả năng truy cập và sự tiện lợi, chứ không phải trên mức độ phù hợp với các câu hỏi nghiên cứu; do đó có thể không đủ khả năng để giải quyết triệt để vấn đề mà câu hỏi nghiên cứu đặt ra. *Thứ ba*, các nhà nghiên cứu thường không sử dụng nhiều kỹ thuật khác nhau một lúc để tăng giá trị hoặc lồng ghép, đối chiếu dữ liệu đã thu thập, điều này có thể dẫn đến giải thích không toàn diện. *Thứ tư*, nhiều nghiên cứu cung cấp rất ít chi tiết về cách thức dữ liệu được thu thập và phân tích (ví dụ, các câu hỏi phỏng vấn đã sử dụng, các tài liệu đã kiểm tra, vị trí người được phỏng vấn trong tổ chức, v.v...), nó có thể làm tăng nghi ngờ về độ tin cậy trong kết luận của họ. *Cuối cùng*, mặc dù có điểm mạnh là một phương pháp nghiên cứu theo chiều dọc, nhưng nhiều nghiên cứu trường hợp lại không quan sát hiện tượng theo phương thức chiều dọc, mà chỉ trình bày một cái nhìn cắt ngang về quy trình phát triển của các hiện tượng trong một khoảng thời gian hạn chế.

⁸ Benbasat, I., Goldstein, D. K., and Mead, M. (1987). "The Case Research Strategy in Studies of Information Systems," *MIS Quarterly* (11:3), 369-386.

Các quyết định quan trọng trong nghiên cứu trường hợp

Nhà nghiên cứu cần phải đưa ra một số quyết định quan trọng khi xem xét tiến hành phương pháp nghiên cứu trường hợp. *Trước hết*, nghiên cứu trường hợp có là phương pháp phù hợp để trả lời cho câu hỏi nghiên cứu? Nghiên cứu trường hợp là phương pháp đặc biệt thích hợp cho các nghiên cứu thăm dò để phát hiện ra những phạm trù có liên quan trong các lĩnh vực nghiên cứu mà việc xây dựng lý thuyết về lĩnh vực này đang ở giai đoạn hình thành. Nó cũng phù hợp cho các nghiên cứu mà kinh nghiệm của những người tham gia và bối cảnh chứa đựng hành vi có vai trò then chốt và cũng thích hợp cho các nghiên cứu nhằm tìm hiểu các quá trình thực tiễn phức tạp (tại sao và làm thế nào một hiện tượng xảy ra) chứ không phải để tìm hiểu mối quan hệ nhân quả. Phương pháp này rất phù hợp để nghiên cứu các quá trình phức tạp của tổ chức có liên quan đến nhiều người tham gia với hàng chuỗi các sự kiện tương tác lẫn nhau, chẳng hạn như dự án cải tổ của tổ chức và việc áp dụng công nghệ ở quy mô lớn.

Thứ hai, các đơn vị phân tích cụ thể trong nghiên cứu trường hợp là gì? Bởi vì nghiên cứu trường hợp có thể kiểm tra đồng thời nhiều đơn vị phân tích, nhà nghiên cứu phải quyết định sẽ nghiên cứu hiện tượng ở cấp độ cá nhân, nhóm, tổ chức, hay kết hợp. Ví dụ, một nghiên cứu về việc ra quyết định nhóm hoặc nghiên cứu về làm việc theo nhóm có thể kết hợp các phạm trù cấp độ cá nhân như sự tham gia của cá nhân trong các hoạt động nhóm với các phạm trù cấp độ nhóm, chẳng hạn như sự gắn kết thành viên nhóm và lãnh đạo nhóm, để thu được những hiểu biết phong phú hơn.

Thứ ba, nhà nghiên cứu có nên sử dụng thiết kế nghiên cứu một trường hợp hay nhiều trường hợp? Thiết kế nghiên cứu một trường hợp thích hợp hơn lúc bắt đầu xây dựng lý thuyết, hay tình huống đang nghiên cứu là duy nhất hoặc đặc biệt hoặc trường hợp mới được nhận biết (ví dụ, các nghiên cứu khoa học trước đây không thể tiếp cận được với tình huống này) hoặc trường hợp này có thể mang đến cơ hội rất tốt để kiểm nghiệm một lý thuyết được xây dựng vững chắc trước đó. Thiết kế nghiên cứu trường hợp thích hợp hơn để thực nghiệm lý thuyết, thiết lập các lập suy luận khái quát, phát triển các giải thích phong phú hơn, nhiều sắc thái hơn về một hiện tượng. Nhà nghiên cứu Yin (1984)⁹ khuyến nghị sử dụng nghiên cứu nhiều trường hợp theo phương thức sao chép mở rộng, nghĩa là xem xét mỗi trường hợp tương tự như một nghiên cứu thực nghiệm, tuân theo các quy tắc khoa học chặt chẽ tương tự như được sử dụng trong nghiên cứu thực chứng.

⁹ Yin, R. K. (2002), *Case Study Research: Design and Methods*. Thousand Oaks, CA: Sage Publications.

Thứ tư, nên lựa chọn địa điểm, đối tượng cụ thể nào trong nghiên cứu trường hợp? Do bản chất phụ thuộc ngữ cảnh của các kết luận thu được từ nghiên cứu trường hợp, lựa chọn địa điểm nghiên cứu phù hợp là một vấn đề đặc biệt quan trọng bởi vì lựa chọn sai có thể dẫn đến những kết luận không khách quan. Nếu mục đích của nghiên cứu là để kiểm tra lý thuyết hay kiểm tra tính khái quát hóa của các suy luận thì các địa điểm khác nhau nên được lựa chọn để tăng tính khác biệt trong các quan sát. Ví dụ, nếu mục tiêu của nghiên cứu là để tìm hiểu quá trình áp dụng công nghệ trong các công ty, một sự kết hợp giữa các công ty cỡ lớn, vừa và nhỏ nên được lựa chọn để kiểm tra xem quá trình áp dụng công nghệ trong các doanh nghiệp với quy mô khác nhau. Lựa chọn địa điểm, đối tượng cụ thể để nghiên cứu không nên dựa trên cơ hội hoặc sự thuận tiện của nhà nghiên cứu, mà nên dựa trên sự phù hợp với các câu hỏi nghiên cứu thông qua quá trình được chọn mẫu.

Thứ năm, các kỹ thuật thu thập dữ liệu nào sẽ được sử dụng trong nghiên cứu trường hợp? Dù rằng cho đến nay phỏng vấn (có thể là mở / phi cấu trúc hay đóng / cấu trúc) là kỹ thuật thu thập dữ liệu phổ biến nhất trong các nghiên cứu trường hợp, tuy vậy dữ liệu thu được từ phỏng vấn có thể được bổ sung hoặc chứng thực bởi các kỹ thuật khác như quan sát trực tiếp (ví dụ như tham gia các cuộc họp quản lý, giao ban và lên chương trình hành động), khảo cứu tài liệu (ví dụ, báo cáo nội bộ, báo cáo thuyết trình, biên bản ghi nhớ, cũng như các tài khoản bên ngoài như báo chí), hồ sơ lưu trữ (ví dụ, sơ đồ tổ chức, hồ sơ tài chính, v.v..), các hiện vật (các công cụ thiết bị kỹ thuật, sản phẩm đầu ra). Hơn nữa, các nhà nghiên cứu nên lồng ghép hoặc củng cố giá trị các dữ liệu thu thập được bằng cách so sánh các phản hồi giữa người trả lời.

Tiến hành nghiên cứu trường hợp

Phần lớn các nghiên cứu trường hợp về bản chất là nghiên cứu diễn giải. Nghiên cứu trường hợp diễn giải là một kỹ thuật quy nạp bằng cách phân tích và tổng hợp một cách hệ thống các thông tin thu thập được từ một hoặc nhiều nguồn, giúp phát hiện các khái niệm và mô hình với mục đích xây dựng lý thuyết mới hoặc mở rộng lý thuyết hiện có. Eisenhardt (1989)¹⁰ đề xuất một “lộ trình” cho việc xây dựng lý thuyết từ các nghiên cứu trường hợp như vậy và một trong những phiên bản đó sẽ được mô tả dưới đây với đôi chút sửa đổi. Đối với nghiên cứu trường hợp thực chứng, một số giai đoạn sau đây có thể cần được sắp xếp lại hoặc sửa đổi lại đôi chút, tuy nhiên quá trình chọn mẫu, thu thập dữ liệu và các kỹ thuật phân tích dữ liệu nói chung vẫn nên giữ nguyên.

¹⁰ Eisenhardt, K. M. (1989). “Building Theories from Case Research,” *Academy of Management Review* (14:4), 532-550.

Xác định câu hỏi nghiên cứu. Giống như bất kỳ dạng nghiên cứu khoa học khác, nghiên cứu trường hợp cũng phải bắt đầu với việc xác định câu hỏi nghiên cứu. Chúng cần phải chú ý về lý thuyết và cả thực tế và sau đó nên hình dung một số phỏng đoán trực quan về các câu trả lời tiềm năng cho những câu hỏi đó hoặc hình dung các phạm trù ban đầu định hướng cho thiết kế nghiên cứu. Trong nghiên cứu trường hợp thực chứng, phạm trù ban đầu được hình thành dựa trên lý thuyết, trong khi đó nghiên cứu trường hợp diễn giải có đặc trưng là không dựa trên một lý thuyết cụ thể và không có các giả thuyết. Sau này, nếu cần thiết những câu hỏi nghiên cứu và phạm trù này có thể được thay đổi trong nghiên cứu trường hợp diễn giải, nhưng không xảy ra trong nghiên cứu trường hợp thực chứng.

Chọn đơn vị nghiên cứu cụ thể (địa điểm và đối tượng nghiên cứu cụ thể). Các nhà nghiên cứu nên sử dụng quá trình "chọn mẫu lý thuyết" (không phải là chọn mẫu ngẫu nhiên) để chọn địa điểm và đối tượng nghiên cứu cụ thể. Trong quá trình này, việc lựa chọn dựa trên những tính toán về mặt lý thuyết chứ không dựa trên các xem xét về thống kê. Những tính toán lý thuyết giúp xác định mục tiêu của nghiên cứu là để kiểm chứng, nhân rộng các nghiên cứu trường hợp trước đó, để mở rộng lý thuyết ban đầu hay để hoàn thiện các mô hình lý thuyết, những mục tiêu này giúp chọn lựa địa bàn, người tham gia nghiên cứu phù hợp. Cần đảm bảo rằng việc lựa chọn những đơn vị này phải phù hợp với bản chất của câu hỏi nghiên cứu, giảm thiểu những yếu tố không liên quan và tối đa hóa sự đa dạng của các biến phụ thuộc. Ví dụ, nếu mục tiêu của nghiên cứu là để kiểm tra tại sao một số công ty sáng tạo hơn so với những công ty khác, các nhà nghiên cứu nên chọn các công ty có quy mô tương tự trong cùng một lĩnh vực để giảm thiểu những sai lệch về ngành nghề kinh doanh và quy mô doanh nghiệp, nhưng đồng thời nên lựa chọn một số công ty có rất nhiều sáng tạo và một công ty có rất ít để tối đa hóa những sai lệch về sự sáng tạo của doanh nghiệp.

Thay vì gọi điện thoại hoặc gửi văn bản tới khách thể nghiên cứu dự kiến, sẽ là tốt hơn nếu nhà nghiên cứu liên lạc với người có thẩm quyền trong tổ chức. Trong khi tiếp xúc ban đầu, các nhà nghiên cứu nên mô tả với khách thể nghiên cứu tính chất và mục đích của đề tài, cơ quan, tổ chức quản lý đề tài, những lợi ích tiềm năng cho khách thể nghiên cứu, cách thức sử dụng các dữ liệu thu thập, những người tham gia thu thập dữ liệu (các nhà nghiên cứu khác, các trợ lý nghiên cứu, vv), những người mà nhà nghiên cứu mong muốn phỏng vấn, thời lượng phỏng vấn. Nhà nghiên cứu còn phải đảm bảo tính bí mật, riêng tư về thông tin cá nhân của người được phỏng vấn.

Chuẩn bị công cụ và trình tự. Bởi vì phương pháp chính để thu thập dữ liệu trong

nghiên cứu trường hợp là các cuộc phỏng vấn. Một trình tự hay kịch bản phỏng vấn nên được hình thành để định hướng quá trình phỏng vấn. Về cơ bản, đó là danh sách các câu hỏi. Câu hỏi có thể mở (phi cấu trúc) hoặc đóng (cấu trúc) hoặc kết hợp cả hai. Trình tự phỏng vấn phải được tuân thủ nghiêm ngặt, người phỏng vấn không được thay đổi thứ tự các câu hỏi hoặc bỏ qua bất kỳ câu hỏi trong quá trình phỏng vấn, tuy nhiên người phỏng vấn có thể đưa thêm một số câu hỏi để khai thác sâu hơn những thông tin của người trả lời mà người hỏi thấy không rõ ràng hoặc thú vị.

Người phỏng vấn phải thể hiện quan điểm trung lập, không định hướng người trả lời đi theo bất kỳ trường phái cụ thể nào, ví dụ đồng ý hoặc không đồng ý với một câu trả lời của người được phỏng vấn. Kỹ thuật phỏng vấn sẽ được thảo luận chi tiết hơn trong chương về khảo sát. Ngoài ra, nên tìm kiếm thêm nguồn dữ liệu, chẳng hạn như tài liệu nội bộ, biên bản ghi nhớ, các báo cáo hàng năm, báo cáo tài chính, bài báo và quan sát trực tiếp để bổ sung, xác nhận dữ liệu phỏng vấn.

Chọn người trả lời. Chọn người trả lời phỏng vấn ở các vị trí công việc khác nhau, các cơ quan, tổ chức khác nhau để có thể có được những quan điểm khác nhau về hiện tượng quan tâm. Một mẫu ngẫu nhiên của người được phỏng vấn là thích hợp nhất, tuy nhiên một mẫu *quả cầu tuyết* (*kỹ thuật mở rộng dần*) cũng chấp nhận được, miễn là có sự đa dạng các quan điểm được thể hiện trong mẫu. Người được phỏng vấn phải được lựa chọn dựa trên sự liên quan cá nhân của họ với hiện tượng đang khảo sát, sự sẵn sàng và khả năng của họ có thể trả lời chính xác và đầy đủ các câu hỏi của nhà nghiên cứu, mà không dựa trên sự thuận tiện của nhà nghiên cứu.

Bắt đầu thu thập dữ liệu. Thông thường, sẽ là một ý tưởng tốt khi dùng các thiết bị điện tử ghi âm lại cuộc phỏng vấn để sử dụng trong tương lai. Tuy nhiên, ghi âm như vậy chỉ phải được thực hiện với sự đồng ý của người được phỏng vấn. Ngay cả khi cuộc phỏng vấn được ghi lại, người phỏng vấn nên bút kí để nắm bắt những thông tin, phản ứng và hành vi quan trọng (ví dụ, ngôn ngữ cơ thể của người trả lời) cũng như ấn tượng cá nhân của nhà nghiên cứu về người trả lời và những bình luận thú vị của họ. Sau mỗi cuộc phỏng vấn, toàn bộ nội dung cuộc phỏng vấn bằng miệng nên được chuyển thể nguyên văn thành dạng tài liệu văn bản để phân tích.

Tiến hành phân tích dữ liệu cho từng trường hợp. Công việc phân tích dữ liệu có thể thực hiện sau hoặc đồng thời với quá trình thu thập dữ liệu. Thu thập và phân tích dữ liệu cùng một lúc có lợi thế là giúp điều chỉnh quá trình thu thập dữ liệu dựa trên kết quả thu

được qua việc phân tích dữ liệu hoặc có thể khai thác sâu thêm các thông tin quan trọng. Phân tích dữ liệu được thực hiện trong hai giai đoạn. Trong giai đoạn đầu (phân tích từng trường hợp), các nhà nghiên cứu nên xem xét một cách riêng biệt các khái niệm mới xuất hiện ở mỗi trường hợp nghiên cứu và mối quan hệ giữa các khái niệm này, sau đó tạo ra một lý thuyết ban đầu về vấn đề nghiên cứu. Các nhà nghiên cứu có thể thu thập dữ liệu theo mong muốn chủ quan kết hợp với việc sử dụng kiến thức và kinh nghiệm cá nhân để trả lời các câu hỏi nghiên cứu cho trường hợp đang nghiên cứu. Ngoài ra, kỹ thuật *mã hoá* (coding), ví dụ như phương pháp *lý thuyết nền* (grounded theory) của Glasser và Strauss (1967), dùng các kỹ thuật như mã hoá mở, mã hoá trực và mã hoá chọn lọc có thể được sử dụng để đạt được một chuỗi luận cứ và kết luận. Những kỹ thuật này được thảo luận chi tiết trong chương sau. Các kỹ thuật *nội sinh* (homegrown), chẳng hạn như phương pháp hiện diện dữ liệu đồ họa (ví dụ sơ đồ mạng) hoặc phân tích trình tự (đối với dữ liệu theo chiều dọc) cũng có thể được sử dụng. Lưu ý rằng không có phương pháp phân tích cố định được áp dụng cho các loại dữ liệu khác nhau trong nghiên cứu trường hợp; các kỹ thuật phân tích dữ liệu có thể được sửa đổi để phù hợp với bản chất của dự án nghiên cứu.

Tiến hành phân tích dữ liệu nhiều trường hợp. Nghiên cứu nhiều trường hợp đòi hỏi phải tiến hành bước phân tích dữ liệu tiếp theo, được gọi là phân tích liên trường hợp. Trong phân tích này, các nhà nghiên cứu cần tìm các khái niệm và các mô hình tương tự giữa các trường hợp khác nhau, bỏ qua sự khác biệt theo ngữ cảnh có thể dẫn đến những kết luận khác biệt. Những mô hình này có thể được sử dụng để chứng thực lý thuyết ban đầu (sau khi phân tích các trường hợp riêng lẻ) hoặc tinh chỉnh nó (bằng cách thêm hoặc bỏ một số khái niệm và mối quan hệ giữa các khái niệm) để phát triển một lý thuyết khái quát và toàn diện hơn. Phân tích này có thể có nhiều hình thức. Ví dụ, các nhà nghiên cứu có thể chọn các phân loại khác nhau (ví dụ, theo quy mô tổ chức, theo lĩnh vực hoạt động, v.v...) và tìm kiếm những tương đồng bên trong nhóm và những khác biệt giữa các nhóm. Ngoài ra, nhà nghiên cứu có thể so sánh đối chiếu từng cặp đôi, rút ra sự giống và khác trong các trường hợp đã nghiên cứu.

Xây dựng và kiểm định các giả thuyết. Dựa trên các khái niệm và các chủ đề xuất hiện sau quá trình phân tích liên trường hợp, các giả thuyết dự kiến sẽ được thiết lập. Những giả thuyết này cần được so sánh lặp đi lặp lại với những dữ liệu đã thu được để xem chúng có phù hợp với các dữ liệu này hay không; nếu không phù hợp, các phạm trù hoặc các mối quan hệ giữa các phạm trù tạo thành giả thuyết cần được điều chỉnh lại.

Ngoài ra, nhà nghiên cứu cũng nên so sánh các phạm trù và giả thuyết mới với những khái niệm và giả thuyết xuất hiện trong các nghiên cứu trước, qua đó giúp nghiên cứu của mình đạt được giá trị nhân quả và giá trị phổ quát cao. Không được loại bỏ những phát hiện có mâu thuẫn, cần tư duy sáng tạo để giải thích các mâu thuẫn đó để hiểu sâu sắc, cặn kẽ lý thuyết mới. Khi việc phân tích lý thuyết và dữ liệu lặp đi lặp lại mà không mang lại những bổ sung, thay đổi so với lý thuyết hiện có, là đã đạt được trạng thái “bão hòa lý thuyết” và quá trình xây dựng lý thuyết hoàn thiện.

Viết báo cáo. Trong văn bản báo cáo, nhà nghiên cứu nên mô tả rõ ràng chi tiết quá trình lấy mẫu, thu thập dữ liệu, phân tích dữ liệu và phát triển các giả thuyết để người đọc có thể đánh giá một cách độc lập tính hợp lý, điểm mạnh, điểm yếu và tính thống nhất của những kết luận được báo cáo. Để đảm bảo các kết quả không bị thiên vị do các định kiến của nhà nghiên cứu, các phương pháp nghiên cứu cần đạt được mức độ cao về sự rõ ràng.

Mô hình mẫu cho nghiên cứu trường hợp điển giải

Cách tốt nhất để tìm hiểu về nghiên cứu trường hợp điển giải là đưa ra một ví dụ minh họa. Nghiên cứu của Eisenhardt (1989)¹¹ về cách thức người lãnh đạo đưa ra quyết định trong môi trường tốc độ cao (high-velocity environments - HVE). Trước khi đọc thông tin tóm tắt về nghiên cứu trong chương này, độc giả nên đọc bản báo cáo gốc được xuất bản tại Tạp chí Quản lý Học thuật (Academy of Management Journal). Trong nghiên cứu này, Eisenhardt khảo sát các nhóm điều hành trong một số tổ chức HVE có quá trình ban hành các quyết định nhanh chóng, trong khi những người trong các tổ chức khác không thể áp dụng được. Eisenhardt tìm hiểu liệu quyết định nhanh hơn có giúp cải thiện hay gây ảnh hưởng xấu đến hoạt động của tổ chức trong môi trường như vậy. HVE được định nghĩa như là một môi trường trong đó nhu cầu, sự cạnh tranh và công nghệ thay đổi rất nhanh chóng nhưng không liên tục, trong khi thông tin trao đổi thường là không chính xác, không có sẵn hoặc đã lỗi thời. Các giả định ban đầu là (1) việc đưa ra quyết định nhanh chóng với thông tin không đầy đủ trong HVE là khó khăn; (2) các quyết định được đưa ra nhanh chóng có thể không có hiệu quả và có thể giảm năng lực hoạt động của tổ chức.

Rà soát các tài liệu đã có về việc ra quyết định của lãnh đạo, Eisenhardt tìm thấy các mô hình khác nhau, tuy vậy không có mô hình cụ thể cho môi trường tốc độ cao. Các tài liệu cho rằng, nhằm hướng tới tính thiết thực, các tổ chức đưa ra quyết định nhanh hơn thì

¹¹ Eisenhardt, K. M. (1989). “Making Fast Strategic Decisions in High-Velocity Environments,” *Academy of Management Journal* (32:3), 543-576.

sẽ ít sử dụng nguồn đầu vào hơn, ít xem xét các lựa chọn thay thế hơn, chỉ đưa ra các phân tích sơ bộ, hạn chế sự tham gia của người có liên quan vào quá trình ra quyết định, tập trung hóa thẩm quyền ra quyết định và hạn chế xung đột nội bộ. Tuy nhiên, Eisenhardt cho rằng những quan điểm này có thể không đúng khi đưa ra quyết định trong môi trường tốc độ cao khi mà các quyết định phải được đưa ra một cách nhanh chóng và với những thông tin không đầy đủ, trong khi đó phải duy trì chất lượng, hiệu quả cao của các quyết định.

Để kiểm tra hiện tượng này, Eisenhardt tiến hành một nghiên cứu quy nạp của tám công ty trong ngành công nghiệp máy tính cá nhân. Các ngành công nghiệp máy tính cá nhân đã trải qua những thay đổi kịch tính trong công nghệ với sự ra đời của hệ điều hành UNIX, RISC và bộ nhớ truy cập ngẫu nhiên 64KB trong những năm 1980, sự cạnh tranh gia tăng với sự xuất hiện của IBM trong các doanh nghiệp máy tính cá nhân và nhu cầu của khách hàng ngày càng tăng với mức tăng trưởng hai con số và do đó thỏa mãn tính chất của môi trường tốc độ cao.

Đây là một thiết kế nhiều trường hợp với logic nhân rộng, theo đó mỗi trường hợp sẽ xác nhận hoặc bác bỏ suy luận từ các trường hợp khác. Các khách thể nghiên cứu được lựa chọn dựa trên khả năng tiếp cận của nhà nghiên cứu, tuy nhiên, tất cả các công ty đều hoạt động trong ngành công nghiệp máy tính cá nhân vận tốc cao trong khu vực Thung lũng Silicon, California. Sự lựa chọn các công ty trong cùng một ngành công nghiệp, cùng khu vực địa lý đã loại trừ tất cả các “nhân tố làm nhiễu” hay những sai số trong các biến phụ thuộc (ở đây là tốc độ ra quyết định và hiệu suất hoạt động).

Nghiên cứu này đã sử dụng một thiết kế đa chiều với nhiều mức độ phân tích: các quyết định (so sánh nhiều quyết định chiến lược khác nhau trong mỗi công ty), các nhóm lãnh đạo (so sánh các nhóm khác nhau chịu trách nhiệm về việc ra các quyết định chiến lược) và công ty (hiệu suất tổng thể của công ty). Dữ liệu được thu thập từ năm nguồn:

- *Phỏng vấn ban đầu với các giám đốc điều hành (CEO)*: Các CEO được hỏi về chiến lược cạnh tranh, năng lực đặc thù, đối thủ cạnh tranh lớn, hiệu suất và quyết định chiến lược quan trọng gần đây của công ty họ điều hành. Dựa trên các cuộc phỏng vấn này, một số quyết định chiến lược trong mỗi công ty được lựa chọn phục vụ việc điều tra tiếp theo. Bốn tiêu chí được sử dụng để lựa chọn các quyết định: (1) các quyết định liên quan đến vị trí chiến lược của công ty, (2) các quyết định đã liên quan đến những đầu tư tài chính lớn, (3) các quyết định liên quan đến nhiều hoạt động khác nhau của công ty, và (4) quyết định đại diện cho quá trình ra hoạch định chiến lược của công ty đó.

- *Phòng vấn với các trưởng phòng*: Mỗi trưởng phòng trả lời 16 câu hỏi mở về chiến lược cạnh tranh, chiến lược hoạt động, thành viên trong các nhóm quản lý chủ chốt, sự thường xuyên và hình thức gặp gỡ của các nhóm này, quá trình diễn hình để đưa ra các quyết định quan trọng và thời gian để đưa ra các quyết định đó. Cuộc phỏng vấn kéo dài từ 1,5 đến 2 giờ và đôi khi kéo dài đến 4 giờ. Để có thể tập trung tìm hiểu các sự kiện thực tế chứ không phải là nhận thức hoặc giải thích của người được hỏi, cách đặt câu hỏi như trong “phòng xử án” được sử dụng, chẳng hạn như điều này xảy ra khi nào? Bạn đã làm gì,.. Các cuộc phỏng vấn được thực hiện bởi hai người và các dữ liệu được củng cố bằng việc kiểm tra trên thực tế bởi người cả hai người phỏng vấn. Tất cả các dữ liệu phỏng vấn đều được ghi âm, tuy nhiên, việc bút ký cũng được thực hiện trong mỗi cuộc phỏng vấn và đã kết thúc với việc ghi lại ấn tượng tổng thể của người phỏng vấn. Sử dụng “nguyên tắc 24 giờ”, các ghi chú chi tiết được ghi lại đầy đủ trong vòng 24 giờ sau cuộc phỏng vấn. Vì vậy, những thông tin hay, ấn tượng của người phỏng vấn dễ dàng được nhớ và ghi lại.

- *Bảng câu hỏi (phiếu khảo sát)*: Các thành viên trong nhóm điều hành tại mỗi công ty hoàn thành một bảng câu hỏi khảo sát để thu thập các dữ liệu định lượng về mức độ xung đột và phân phối quyền lực trong công ty của họ.

- *Dữ liệu thứ cấp*: Các báo cáo ngành công nghiệp máy tính và các tài liệu nội bộ như nhân sự của các nhóm điều hành (chịu trách nhiệm về các quyết định chiến lược), hoạt động tài chính của các công ty,... đã được kiểm tra.

- *Quan sát cá nhân*: Cuối cùng, các nhà nghiên cứu tham dự một phiên họp chiến lược 1 ngày và một cuộc họp điều hành hàng tuần tại hai công ty trong mẫu.

Quá trình phân tích dữ liệu đã kết hợp các kỹ thuật định lượng và định tính. Dữ liệu định lượng về xung đột và quyền lực được phân tích để tìm ra các mô hình chung xuyên suốt các công ty và các quyết định. Các dữ liệu phỏng vấn định tính được kết hợp để tìm hiểu quá trình ra quyết định, bằng cách sử dụng các đặc điểm cá biệt (ví dụ, sự thiếu kiên nhẫn) được đề cập bởi nhiều hơn một nhà điều hành. Đối với phân tích trong mỗi trường hợp, quá trình ra quyết định chiến lược được mô tả bằng cách kết hợp thông tin của người điều hành theo trình tự thời gian. Đối với phân tích tổng hợp nhiều trường hợp, từng cặp đôi công ty được so sánh để tìm ra điểm giống và khác nhau, được phân loại theo các biến số cần quan tâm như tốc độ ra quyết định và hiệu quả hoạt động của công ty. Dựa trên phân tích này, các phạm trù và luận điểm dự kiến được hình thành theo cách quy nạp. Mỗi trường hợp ra quyết định được xem xét lại để xác nhận các mối quan hệ dự kiến. Bằng

cách so sánh những kết quả của quá trình khảo cứu tài liệu với các luận điểm được suy luận, nhà nghiên cứu đã có được những hiểu biết mới về các trường hợp. Cuối cùng, lý thuyết về quá trình ra quyết định chiến lược của các công ty trong môi trường tốc độ cao được hình thành một cách quy nạp thông qua tổng hợp các luận điểm đã củng cố.

Các suy luận xuất phát từ nghiên cứu nhiều trường hợp như thế này đã mâu thuẫn với một số mô hình ra quyết định được đề xuất trong các nghiên cứu trước đó. *Thứ nhất*, không như dự kiến của các tài liệu trước đó, những người ra quyết định nhanh trong môi trường tốc độ cao đã sử dụng nhiều nguồn thông tin hơn. Tuy nhiên, những người ra quyết định sử dụng nhiều thông tin *hữu ích* (một khám phá mới chưa được đề cập trong các nghiên cứu trước), nó giúp họ xác định và xử lý với các vấn đề và các thay đổi một cách nhanh chóng hơn. *Thứ hai*, những người đưa ra quyết định nhanh đã xem xét nhiều hơn những lựa chọn thay thế. Tuy nhiên, họ xem xét cùng một lúc nhiều lựa chọn thay thế, trong khi đó những người ra quyết định chậm hơn lại xem xét số lượng ít hơn các lựa chọn thay thế và theo một cách tuần tự. *Thứ ba*, những người ra quyết định nhanh không tập trung hóa quá trình ra quyết định hoặc hạn chế việc tham gia của những người có liên quan, như các tài liệu trước đó đã đề cập. Thay vào đó, các công ty này chia quá trình ra quyết định thành hai giai đoạn, trong đó ở giai đoạn đầu tiên, nhiều nhân viên tư vấn giàu kinh nghiệm được đề nghị đưa ra ý kiến đóng góp; tiếp theo giai đoạn thứ hai là quá trình so sánh và lựa chọn một cách nhanh chóng các quyết định. *Thứ tư*, những người ra quyết định nhanh không có ít xung đột hơn như dự báo từ các nghiên cứu trước, nhưng họ lại sử dụng các kỹ thuật giải quyết xung đột tốt hơn để giảm mâu thuẫn và cải thiện tốc độ ra quyết định. *Cuối cùng*, những người đưa ra quyết định nhanh tận dụng tối đa mặt mạnh trong công ty của họ để đẩy nhanh tiến trình ra các quyết định quan trọng.

Ví dụ về nghiên cứu trường hợp thực chứng

Nghiên cứu trường hợp cũng có thể được sử dụng cho các nghiên cứu thực chứng để kiểm định lý thuyết hay giả thuyết. Những nghiên cứu như vậy là không nhiều. Markus (1983)¹² cung cấp một minh họa trong nghiên cứu của cô về việc áp dụng công nghệ tại Công ty Tam giác vàng (Golden Triangle Company - GTC). Mục tiêu của nghiên cứu này là tìm hiểu lý do tại sao một hệ thống thông tin tài chính mới (FIS) áp dụng để nâng cao hiệu quả công tác kế toán tại GTC được các kế toán viên tại hội sở công ty GTC ủng hộ, nhưng lại bị các kế toán viên cấp phòng tại các chi nhánh của GTC phản đối. Với sự độ

¹² Markus, M. L. (1983). "Power, Politics, and MIS Implementation," *Communications of the ACM* (26:6), 430-444.

đáo của vấn đề nghiên cứu, đây là nghiên cứu một trường hợp.

Để khám phá những lý do đằng sau sự phản đối sử dụng FIS, Markus đưa ra ba lý thuyết giải thích khác nhau: (1) lý thuyết *hệ thống quyết định*: sự phản đối là bởi các yếu tố liên quan đến hệ thống không phù hợp, chẳng hạn như thiếu sót kỹ thuật, thiết kế kém tiện dụng hoặc thiếu thân thiện với người sử dụng, (2) lý thuyết *người sử dụng quyết định*: sự phản đối là bởi các yếu tố bản thân người sử dụng, chẳng hạn như nhận thức hoặc các đặc điểm tính cách của các kế toán viên không phù hợp với việc sử dụng hệ thống; và (3) lý thuyết *tương tác*: sự phản đối là bởi các yếu tố bản thân người sử dụng và của hệ thống. Cụ thể, lý thuyết tương tác gợi ý rằng FIS nảy sinh bởi có sự phân bổ lại quyền lợi trong nội bộ tổ chức, những kế toán bị mất vị trí và thẩm quyền trong tổ chức khi áp dụng FIS phản đối hệ thống, trong khi những người có thêm quyền lợi lại ủng hộ.

Để kiểm tra ba lý thuyết này, Markus đã dự đoán các giải pháp cho mỗi giả thuyết và phân tích mức độ thành công của các giải pháp đó tại GTC. Ví dụ, lý thuyết *hệ thống quyết định* gợi ý rằng sự phản đối của người sử dụng gây ra bởi vì hệ thống không phù hợp, vì vậy sửa chữa các vấn đề kỹ thuật của hệ thống sẽ loại bỏ sự phản đối này. Máy tính đang chạy hệ thống FIS sau đó được nâng cấp với một hệ điều hành mạnh hơn, xử lý trực tuyến (so với ban đầu là việc xử lý hàng loạt, đã trì hoãn yêu cầu xử lý ngay lập tức các thông tin kế toán) và một phần mềm đơn giản hơn được áp dụng để tạo tài khoản mới cho các nhà quản lý. Một năm sau khi những thay đổi này được thực hiện, người dùng vẫn phản đối hệ thống và cho rằng nó nên được thay thế. Do đó, lý thuyết hệ thống quyết định đã bị loại bỏ.

Lý thuyết *người dùng quyết định* dự đoán rằng thay thế những cá nhân phản đối hoặc phân công họ làm việc cùng với những người dùng ít tranh cãi sẽ làm giảm sự phản đối của họ đối với FIS. Sau đó, GTC bắt đầu luân phiên công việc và thực hiện chính sách di động, điều chuyển kế toán viên trong những đơn vị có phản đối. Tuy nhiên, sức phản đối không những tiếp tục mà một số trường hợp tăng lên! Một ví dụ cụ thể, một kế toán là một trong những nhà thiết kế và những người ủng hộ hệ thống FIS khi ông làm việc tại cấp kế toán tổng hợp, đã bắt đầu phản đối hệ thống sau khi ông được chuyển xuống cấp phòng. Thất bại trong việc chứng minh các giả thuyết của lý thuyết người dùng quyết định dẫn đến việc bác bỏ lý thuyết này.

Cuối cùng, lý thuyết tương tác dự báo rằng dù có tác động vào hệ thống hay vào người dùng (ví dụ, giáo dục người sử dụng hoặc sử dụng chính sách luân chuyển nhân viên) đều không làm giảm sự phản đối trừ khi vấn đề thiếu công bằng về quyền lợi và việc

phân phối lại được thực hiện. Trước khi thực hiện FIS, kế toán viên cấp phòng của GTC cảm thấy rằng họ sở hữu tất cả các dữ liệu liên quan đến hoạt động kế toán của phòng họ làm việc. Họ đã lưu dữ liệu này trong những cuốn sổ cái thủ công dày, kiểm soát được sự truy cập của những người khác và có thể điều chỉnh các sự cố kế toán bất thường trước khi công bố các báo cáo. Kế toán tổng hợp dựa nhiều vào kế toán ở các phòng để truy cập vào các dữ liệu để làm các báo cáo tổng hợp của công ty. Bởi vì hệ thống FIS tự động thu thập tất cả các dữ liệu tại nguồn và củng cố chúng thành một cơ sở dữ liệu duy nhất của công ty, nó loại bỏ sự cần thiết của các kế toán cấp phòng, làm mất đi sự kiểm soát và quyền của họ đối với các số liệu kế toán của phòng; làm cho công việc của họ phần nào không quan trọng. Kế toán tổng hợp bây giờ có thể truy cập và truy vấn trực tiếp cơ sở dữ liệu của các phòng mà không cần phải thông qua kế toán ở đó; việc phân tích và đánh giá hiệu suất công việc của từng phòng và trực tiếp báo cáo các dấu hiệu khác thường tới Ban giám đốc, điều này một lần nữa dẫn đến sự xói mòn quyền lực của các phòng. Mặc dù Markus không thực nghiệm kiểm tra lý thuyết này, nhưng qua quan sát của mình về phân bổ lại quyền lực của tổ chức, cùng với sự loại trừ của hai lý thuyết trước đã dẫn đến sự giải thích của lý thuyết tương tác là đúng.

So sánh với nghiên cứu truyền thống

Nghiên cứu trường hợp thực chứng nhằm mục đích kiểm nghiệm giả thuyết thường bị chỉ trích bởi các nhà nghiên cứu khoa học tự nhiên vì thiếu những quan sát và suy luận đối chứng; các kết quả thiếu tính thống nhất và tính phổ quát - những nguyên tắc truyền thống trong nghiên cứu thực chứng. Tuy nhiên, những lời chỉ trích có thể được khắc phục thông qua các thiết kế nghiên cứu trường hợp phù hợp. *Thứ nhất*, vấn đề quan sát đối chứng đề cập đến những khó khăn để đạt được sự đối chứng thống kê hay đối chứng thực nghiệm trong nghiên cứu trường hợp. Trên thực tế, các nhà nghiên cứu trường hợp vẫn có thể cải thiện vấn đề này bằng cách sử dụng “đối chứng tự nhiên”. Kỹ thuật đối chứng tự nhiên này trong nghiên cứu của Markus (1983) chính là người kế toán tổng hợp. Ban đầu, anh này là một trong những ủng hộ hệ thống FIS, nhưng sau đó bắt đầu phản đối khi anh ta chuyển đến cấp phòng. Trong trường hợp này, sự thay đổi hành vi của anh này có thể là do bị chuyển vị trí.

Tuy nhiên, những đối chứng tự nhiên như vậy khó có thể được dự đoán trước và các nhà nghiên cứu trường hợp có thể không để ý, trừ khi các nhà nghiên cứu chủ động tìm kiếm các đối chứng như vậy. Một cách ngẫu nhiên, đối chứng tự nhiên cũng được sử dụng trong các ngành khoa học tự nhiên như thiên văn học, địa chất học và sinh học, chẳng hạn

như việc chờ sao chổi chuẩn bị băng qua gần trái đất để đưa ra các suy luận về sao chổi và thành phần của nó.

Thứ hai, vấn đề suy luận đối chứng liên quan đến việc thiếu những bằng chứng định lượng thuyết phục để hỗ trợ các kết luận khi dữ liệu trong nghiên cứu trường hợp chủ yếu là định tính. Mặc dù thiếu các dữ liệu định lượng để kiểm tra giả thuyết, suy luận đối chứng vẫn có thể có được trong nghiên cứu trường hợp bằng cách tạo ra các dự đoán hành vi dựa trên những xem xét lý thuyết và xem xét thử nghiệm về những dự đoán đó trong một khoảng thời gian nhất định. Markus sử dụng chiến lược này trong nghiên cứu của mình bằng cách tạo ra ba giả thuyết thay thế về sự phản đối của người sử dụng và loại bỏ hai trong số những dự đoán khi chúng không phù hợp với các hành vi được quan sát trên thực tế. Trong trường hợp này, các giả thuyết được kiểm nghiệm bằng cách sử dụng các luận điểm logic mà không sử dụng các kiểm tra toán học, điều này cũng có giá trị tương đương với các suy luận thống kê, bởi toán học là một tập hợp con của logic học.

Thứ ba, tính lặp lại (replicability) đề cập đến những khó khăn trong việc nghiên cứu cùng một vấn đề nhưng ở khách thể nghiên cứu khác có những đặc điểm đặc thù. Tuy nhiên, lấy việc sử dụng ba lý thuyết của Markus như là một minh họa, một nhà nghiên cứu khác có thể kiểm nghiệm cùng một lý thuyết đối với một khách thể nghiên cứu (trường hợp nghiên cứu) khác, dựa trên đặc điểm đặc thù của khách thể nghiên cứu mới và ba giải thích khác nhau có thể xuất hiện, sau đó các giải thích này sẽ được kiểm tra. Nói cách khác, có thể tiến hành lặp lại các kết luận của nghiên cứu trường hợp, cho dù khách thể nghiên cứu hoặc bối cảnh của trường hợp mới có thể không đồng nhất với khách thể và bối cảnh của trường hợp được nghiên cứu trước đó.

Thứ tư, nghiên cứu trường hợp có xu hướng khám phá một hiện tượng riêng biệt và không thể nhân rộng, khó để khái quát hóa cho các hiện tượng khác. Tuy vậy, tính phổ quát vẫn có thể được thiết lập thông qua các nghiên cứu bổ sung. Các nghiên cứu trường hợp bổ sung này nếu được thực hiện một cách liên tục, dù trong các ngữ cảnh khác nhau với các giả thuyết khác nhau, nhưng mang đến những kết quả nghiên cứu có liên quan, có hệ thống thì chúng có thể được tổng hợp để hình thành nên tính phổ quát.

Cuối cùng, triết gia người Anh Karl Popper đã mô tả bốn yêu cầu của lý thuyết khoa học: (1) lý thuyết có thể chứng nghiệm là sai lầm, (2) lý thuyết nhất quán về lập luận (3) lý thuyết phải có đủ khả năng tiên đoán, (4) lý thuyết phải đưa ra những lời giải thích thuyết phục hơn so với các lý thuyết phủ định. Trong nghiên cứu trường hợp, khả năng đáp ứng

ba yêu cầu đầu tiên có thể được cải thiện bằng cách tăng bậc tự do của các kết quả nghiên cứu, chẳng hạn như bằng cách tăng số lượng địa bàn nghiên cứu, trường hợp nghiên cứu, tăng số lượng các giả thuyết thay thế và đa dạng hóa các kỹ thuật, mức độ phân tích dữ liệu. Điều này được thực hiện trong nghiên cứu của Markus bằng cách kiểm tra hành vi của nhiều nhóm (các kế toán phòng và các kế toán tổng hợp) và cung cấp nhiều (ba) giả thuyết thay thế. Điều kiện thứ tư của Popper đã đạt được trong nghiên cứu này khi một giả thuyết được chứng minh là phù hợp hơn với những bằng chứng đã quan sát và thuyết phục hơn so với hai giả thuyết thay thế.

Chương 12

NGHIÊN CỨU DIỄN GIẢI

Chương trước giới thiệu về một loại nghiên cứu trường hợp diễn giải. Chương này sẽ khám phá các loại nghiên cứu diễn giải khác. Như đã biết, trong nghiên cứu thực chứng hay diễn dịch, các phương pháp như thử nghiệm trong phòng thí nghiệm và nghiên cứu khảo sát được ưu tiên sử dụng để kiểm nghiệm lý thuyết (hay các giả thuyết), trong khi đó trong nghiên cứu diễn giải hoặc quy nạp, các phương pháp như nghiên cứu hành động và dân tộc học phù hợp cho xây dựng lý thuyết. Không giống như phương pháp thực chứng, nơi mà nhà nghiên cứu bắt đầu với một lý thuyết và kiểm tra các định đề hay luận điểm lý thuyết bằng cách sử dụng các dữ liệu thực nghiệm, ở phương pháp diễn giải, nhà nghiên cứu bắt đầu với dữ liệu và cố gắng để hình thành một lý thuyết về hiện tượng quan tâm từ các dữ liệu quan sát.

Thuật ngữ "*nghiên cứu diễn giải*" thường được gán và đồng nghĩa với "*nghiên cứu định tính*", mặc dù hai khái niệm khác nhau cơ bản. Nghiên cứu diễn giải là một mô hình nghiên cứu (xem Chương 3) dựa trên giả định rằng thực tế xã hội không tồn tại độc lập hay khách quan, mà được định hình bởi kinh nghiệm của con người và bối cảnh xã hội (bản thể luận), do đó sẽ là tốt nhất nếu được nghiên cứu trong bối cảnh kinh tế - lịch sử bằng cách kết hợp những giải thích chủ quan của những người tham gia khác nhau (nhận thức luận). Bởi các nhà nghiên cứu diễn giải nhận định thực tiễn xã hội gắn bó chặt chẽ, không thể tách khỏi môi trường và bối cảnh xã hội; họ “giải thích” (interpret) thực tiễn thông qua quá trình “tạo ý nghĩa” (sense-making) chứ không phải là một quá trình kiểm tra giả thuyết. Điều này là trái ngược với mô thức thực chứng (positivist paradigm) hoặc mô thức chức năng luận (functionalist paradigm), cho rằng thực tế là tương đối độc lập với bối cảnh, có thể được tách biệt khỏi bối cảnh của nó, có thể được nghiên cứu bằng các công thức toán học, sử dụng các kỹ thuật khách quan được tiêu chuẩn hóa. Việc một nhà nghiên cứu tiến hành nghiên cứu thực chứng hay nghiên cứu diễn giải phụ thuộc vào việc cân nhắc kỹ lưỡng bản chất của hiện tượng nghiên cứu và mô thức nào là tốt nhất để nghiên cứu nó.

Trong khi đó, nghiên cứu định lượng hay nghiên cứu định tính đề cập đến tính chất của dữ liệu cần thu thập và cách thức thu thập, phân tích chúng. Nghiên cứu định tính dựa chủ yếu vào các dữ liệu “không phải con số”, chẳng hạn thông tin từ các cuộc phỏng vấn

và quan sát, trái ngược với nghiên cứu định lượng sử dụng dữ liệu số như điểm số và các số liệu khác. Do đó, nghiên cứu định tính không buộc phải tuân theo các thủ tục thống kê như phân tích hồi quy, nhưng lại được *mã hoá* bằng các kỹ thuật định tính như kỹ thuật phân tích nội dung. Đôi khi, dữ liệu định lượng sau khi mã hoá được sắp xếp một cách định tính ví như độ thường xuyên của các *mã hoá*, tuy nhiên dữ liệu này không được phân tích theo cách thống kê số học. Nhiều nhà nghiên cứu điển giải thuần túy không ủng hộ phương pháp *mã hoá*, coi đó là vô ích bởi vì không bao giờ có thể đạt được sự đồng thuận hoặc tính khách quan về một hiện tượng xã hội khi hiện tượng này về bản chất tồn tại một cách chủ quan.

Mặc dù nghiên cứu điển giải thường dựa nhiều vào dữ liệu định tính, dữ liệu định lượng có thể được bổ sung để giúp hiểu chính xác hơn và rõ hơn về hiện tượng quan tâm so với việc chỉ sử dụng các dữ liệu định tính. Ví dụ, Eisenhardt (1989), trong nghiên cứu điển giải về ra quyết định trong các công ty vận tốc cao (được thảo luận trong chương trước về nghiên cứu trường hợp), đã thu thập các dữ liệu số về khoảng thời gian để mỗi doanh nghiệp để đưa ra các quyết định chiến lược (dao động từ 1,5 tháng đến 18 tháng), có bao nhiêu lựa chọn thay thế được xem xét cho mỗi quyết định và có bao nhiêu người được hỏi về xung đột trong tổ chức. Số dữ liệu đó đã giúp Eisenhardt phân định rõ ràng sự khác biệt giữa các công ty ra quyết định tốc độ cao và các công ty ra quyết định tốc độ thấp, mà không dựa hoàn toàn vào nhận thức chủ quan của người trả lời, sau đó nó cho phép Eisenhardt kiểm tra, so sánh số lựa chọn thay thế được xem xét và mức độ xung đột trong các công ty có tốc độ cao và tốc độ thấp. Nghiên cứu điển giải nên cố gắng thu thập cả dữ liệu định tính và định lượng liên quan đến hiện tượng mà họ quan tâm và nghiên cứu thực chứng cũng nên như vậy. Việc sử dụng kết hợp dữ liệu định tính và định lượng như vậy, thường được gọi là "*thiết kế kết hợp*", nó sẽ mang đến những hiểu biết độc đáo và được đánh giá cao trong cộng đồng khoa học.

Phân biệt với nghiên cứu thực chứng

Ngoài sự khác biệt cơ bản về mô thức trong các giả định bản thể luận và nhận thức luận được thảo luận ở trên, nghiên cứu điển giải và thực chứng còn khác nhau ở một số góc độ khác. *Trước hết*, nghiên cứu điển giải sử dụng chiến lược *lấy mẫu lý thuyết*. Đó là việc những nơi, những người trả lời hoặc các trường hợp nghiên cứu được lựa chọn dựa trên những xem xét về mặt lý thuyết, chẳng hạn như liệu chúng có phù hợp với hiện tượng được nghiên cứu (ví dụ, vấn đề phát triển bền vững chỉ có thể được nghiên cứu trong các tổ chức đã và đang thực hiện chiến lược phát triển bền vững); liệu rằng chúng có các đặc

điểm thích hợp nhất cho nghiên cứu (ví dụ, một nghiên cứu về các yếu tố làm nên sự sáng tạo trong các doanh nghiệp nên bao gồm một số công ty có khả năng sáng tạo cao và một số là thấp, để rút ra độ tương phản giữa các công ty),... Ngược lại, nghiên cứu thực chứng sử dụng *lấy mẫu ngẫu nhiên* (hoặc một biến thể của kỹ thuật này), các trường hợp được lựa chọn ngẫu nhiên từ tổng thể khách thể nghiên cứu nhằm mục đích khái quát hóa. Như vậy, các mẫu tiện lợi và mẫu nhỏ là chấp nhận được trong nghiên cứu diễn giải miễn là chúng phù hợp với tính chất và mục đích của các nghiên cứu này, nhưng các mẫu đó lại không được chấp nhận trong nghiên cứu thực chứng.

Thứ hai, vai trò của nhà nghiên cứu nhận được sự quan tâm đặc biệt trong các nghiên cứu diễn giải. Trong một số phương pháp như nghiên cứu dân tộc học, nghiên cứu hành động và quan sát tham dự, nhà nghiên cứu được coi là một phần của hiện tượng xã hội, có vai trò cụ thể và sự tham gia của nhà nghiên cứu vào quá trình nghiên cứu phải được thực hiện rõ ràng trong quá trình phân tích dữ liệu. Trong các phương pháp khác, chẳng hạn như nghiên cứu trường hợp, nhà nghiên cứu phải có lập trường “trung lập” hoặc quan điểm không thiên vị trong suốt quá trình thu thập và phân tích dữ liệu nhằm đảm bảo chắc chắn rằng những thành kiến hay giả định cá nhân của người nghiên cứu không làm sai lệch bản chất của những suy luận chủ quan xuất phát từ nghiên cứu giải thích. Trong nghiên cứu thực chứng, nhà nghiên cứu được coi là bên ngoài và độc lập với bối cảnh nghiên cứu, vậy nên thiên vị cá nhân khó ảnh hưởng hơn tới hoạt động thu thập và phân tích dữ liệu.

Thứ ba, phân tích diễn giải hướng đến tổng thể và theo ngữ cảnh, chứ không phải là theo từng phần nhỏ và biệt lập. Các giải thích diễn giải có xu hướng tập trung vào ngôn ngữ, dấu hiệu và ý nghĩa từ quan điểm của những người tham gia có liên quan đến hiện tượng xã hội, trái ngược với kỹ thuật thống kê được sử dụng rất nhiều trong nghiên cứu thực chứng. Giá trị của phương pháp nghiên cứu diễn giải được đánh giá dựa trên cách thức tiếp cận có hệ thống và minh bạch để thu thập và phân tích dữ liệu, chứ không phải là các tiêu chuẩn thống kê về các giá trị phạm trù.

Cuối cùng, hoạt động thu thập, phân tích dữ liệu có thể tiến hành đồng thời và lặp đi lặp lại trong nghiên cứu giải thích. Ví dụ, nhà nghiên cứu có thể tiến hành một cuộc phỏng vấn và phân tích trước khi tiến hành cuộc phỏng vấn tiếp theo. Việc tiến hành phân tích đồng thời giúp nhà nghiên cứu nhận diện các lỗi tiềm năng trong quá trình phỏng vấn và điều chỉnh trình tự phỏng vấn sao cho thu thập được những thông tin chính xác hơn về những hiện tượng quan tâm. Nhà nghiên cứu thậm chí có thể thay đổi câu hỏi nghiên cứu ban đầu, nếu họ nhận thấy câu hỏi nghiên cứu ban đầu không có khả năng để tạo ra những

tri thức mới và hữu ích. Đây là một lợi ích rất đáng khai thác nhưng thường không được các nhà nghiên cứu diễn giải tận dụng. Lợi thế này không có trong nghiên cứu thực chứng khi mà dự án nghiên cứu không thể được sửa đổi hoặc thay đổi khi đã bắt đầu thu thập dữ liệu trừ khi bắt đầu làm lại toàn bộ dự án.

Lợi ích và thách thức của nghiên cứu diễn giải

Nghiên cứu diễn giải có một số lợi thế đặc thù. *Trước tiên*, chúng rất phù hợp để khám phá những lý do ẩn đằng sau một quá trình xã hội phức tạp, tìm hiểu các quan hệ hữu cơ đa chiều, chẳng hạn như lợi ích nhóm trong tổ chức hoặc vấn đề chính trị trong cơ quan. Để tìm hiểu những vấn đề như vậy, bằng chứng định lượng có thể bị sai lệch, không chính xác hoặc khó khăn để có được. *Thứ hai*, chúng thường hữu ích cho việc xây dựng lý thuyết trong những lĩnh vực mới mà trước đó chưa có lý thuyết hoặc lý thuyết không đầy đủ. *Thứ ba*, chúng cũng thích hợp cho việc nghiên cứu những sự kiện hay quy trình riêng biệt, mang phong cách riêng. *Thứ tư*, nghiên cứu diễn giải cũng có thể giúp phát hiện ra các câu hỏi nghiên cứu thích hợp và thú vị cũng như các vấn đề khoa học cho các nghiên cứu tiếp theo.

Đồng thời, nghiên cứu diễn giải cũng đối mặt với một số thách thức. *Trước hết*, nghiên cứu này thường đòi hỏi đầu tư thời gian, nguồn lực và công sức nhiều hơn, chuyên sâu hơn so với nghiên cứu thực chứng trong việc thu thập và phân tích dữ liệu. Quá ít dữ liệu có thể dẫn đến những giả định sai lầm ấu trĩ, trong khi quá nhiều dữ liệu có thể sẽ không được xử lý hiệu quả.

Thứ hai, nghiên cứu diễn giải đòi hỏi nhà nghiên cứu phải được đào tạo kỹ lưỡng để có khả năng phát hiện và giải thích các hiện tượng xã hội phức tạp từ quan điểm của những người tham gia, phối kết hợp các quan điểm đa dạng và đa chiều của họ, mà không được lòng ghép hay bị ảnh hưởng bởi những thành kiến hay giả định cá nhân của mình khi đưa ra các suy luận khoa học.

Thứ ba, những người tham gia hoặc các nguồn dữ liệu có thể không đồng đều về độ tin cậy, mức độ am hiểu hay sự thiên vị về hiện tượng quan tâm hoặc họ có thể biết nhưng lại không tiết lộ những vấn đề nhạy cảm. Tất cả đều có thể dẫn đến hiểu lầm hoặc ấn tượng sai của nhà nghiên cứu. Thiếu sự tin tưởng của người tham gia đối với nhà nghiên cứu cũng có thể là một rào cản tới sự nhiệt tình và trung thực của người tham gia, trong khi việc xây dựng niềm tin như vậy cần phải có thời gian. Nhiệm vụ của nhà nghiên cứu giải thích là phải “nhìn xuyên màn khói” (những vấn đề ẩn hoặc thành kiến) để hiểu được bản chất thực sự của vấn đề.

Thứ tư, bị ràng buộc chặt chẽ bởi bối cảnh đặc thù, các kết luận rút ra từ nghiên cứu diễn giải, không dễ để nhân rộng hoặc khái quát hóa. Cuối cùng, nghiên cứu diễn giải đôi khi không thể trả lời các câu hỏi nghiên cứu hoặc có thể không tiên đoán được các hành vi trong tương lai.

Đặc điểm của nghiên cứu diễn giải

Tất cả các nghiên cứu diễn giải phải tuân theo một tập hợp các nguyên tắc chung, như được mô tả dưới đây.

Tuân thủ quy luật tự nhiên. Các hiện tượng xã hội phải được nghiên cứu trong bối cảnh tự nhiên của chúng. Bởi vì nghiên cứu diễn giải giả định rằng hiện tượng xã hội nằm bên trong và không thể tách biệt khỏi bối cảnh xã hội của chúng nên giải thích các hiện tượng như vậy phải đặt trong bối cảnh lịch sử - xã hội đó. Điều này yêu cầu khi tìm kiếm lời giải thích cho một hiện tượng quan tâm, các biến phải được quan sát và xem xét theo ngữ cảnh, mặc dù điều này có thể hạn chế tính phổ quát của các suy luận.

Nhà nghiên cứu là một công cụ. Các nhà nghiên cứu nên thấu hiểu bối cảnh xã hội mà họ đang nghiên cứu và họ được coi là một phần của công cụ thu thập dữ liệu; trong đó họ phải sử dụng các kỹ năng quan sát của mình, tạo dựng sự tin tưởng đối với những người tham gia và sử dụng khả năng trích xuất các thông tin chính xác từ những gì đã quan sát. Hơn nữa, những hiểu biết cá nhân, kiến thức và kinh nghiệm về bối cảnh xã hội là rất quan trọng để đưa ra những giải thích chính xác về các hiện tượng quan tâm. Đồng thời, nhà nghiên cứu phải có nhận thức đầy đủ về những thành kiến và giả định cá nhân của mình, không để cho những định kiến như vậy ảnh hưởng đến việc trình bày một bức tranh công bằng và chính xác của hiện tượng này.

Phân tích diễn giải. Quan sát phải được diễn giải bằng quan điểm của những người tham gia mà không tách rời trong bối cảnh xã hội. Diễn giải phải xảy ra ở hai cấp độ. Mức đầu tiên là phải xem xét hiện tượng này từ quan điểm chủ quan của những người tham gia. Cấp độ thứ hai là hiểu được ý nghĩa những kinh nghiệm mà người tham gia đã trải qua để cung cấp một “mô tả sâu” (thick description) hoặc một câu chuyện tự sự phong phú về hiện tượng quan tâm, đưa ra lý do tại sao người tham gia lại hành động theo cách mà họ đã làm.

Sử dụng ngôn ngữ biểu cảm. Khám phá và phân tích ngôn ngữ bằng lời nói và cả những ngôn ngữ biểu cảm không lời của người tham gia là thành phần không thể tách rời của phân tích diễn giải. Người nghiên cứu phải đảm bảo rằng câu chuyện được phản ánh qua ánh mắt của một con người, không phải là một cái máy; phải thể hiện những cảm xúc

và kinh nghiệm của người đó, để người đọc có thể hiểu và hình dung về người đó. Việc sử dụng tranh ảnh, lời nói ẩn dụ, ngôn ngữ châm biếm và các minh họa khác là rất phổ biến trong phân tích diễn giải.

Theo dõi qua thời gian. Nghiên cứu diễn giải thường không quan tâm đến việc tìm kiếm các câu trả lời cụ thể, mà đi tìm hiểu hoặc “tìm hiểu ý nghĩa” một quá trình xã hội năng động tiếp diễn trong khoảng thời gian nhất định. Do đó, nghiên cứu này đòi hỏi sự tham gia của nhà nghiên cứu với khách thể nghiên cứu trong một khoảng thời gian tương đối dài để nắm bắt toàn bộ sự phát triển của hiện tượng quan tâm.

Chu trình diễn ngôn. Giải thích diễn giải là một quá trình lặp lại của các hoạt động nghiên cứu và các hiện tượng xã hội để hài hòa những mâu thuẫn giữa chúng từ đó xây dựng những lý thuyết phù hợp với quan điểm chủ quan và trải nghiệm cá nhân của những người tham gia. Sự lặp đi lặp lại giữa hiểu biết về một hiện tượng và các quan sát phải tiếp tục cho đến khi đạt được sự “bão hòa lý thuyết”, khi đó dù có tiến hành thêm bất kỳ sự lặp lại nào nữa cũng không mang lại cái nhìn tri thức mới và sâu sắc về hiện tượng quan tâm.

Thu thập dữ liệu diễn giải

Dữ liệu được thu thập trong nghiên cứu diễn giải bằng nhiều kỹ thuật khác nhau. Các kỹ thuật được sử dụng thường xuyên nhất là **phỏng vấn** (trực diện, qua điện thoại hoặc hội thảo nhóm nhỏ). Các hình thức và chiến lược phỏng vấn được thảo luận chi tiết trong chương trước về nghiên cứu khảo sát. Kỹ thuật thứ hai là **quan sát**. Kỹ thuật quan sát bao gồm *quan sát trực tiếp (direct observation)*, khi các nhà nghiên cứu là một người quan sát bên ngoài, trung tính và thụ động và không tham gia vào hiện tượng đang quan sát (như trong nghiên cứu trường hợp) và *quan sát tham dự (participant observation)*, khi các nhà nghiên cứu là người nhiệt tình tham gia hoạt động vào trong các sự kiện đang được nghiên cứu và những ảnh hưởng tạo ra bởi sự hiện diện của nhà nghiên cứu sẽ được nhận diện và phân tích (như trong nghiên cứu hành động). Kỹ thuật thứ ba là *điều tra văn bản (documentation)*. Đây là việc phân tích, tổng hợp các tài liệu nội bộ và tài liệu khác có liên quan, chẳng hạn như các bản ghi nhớ, thư điện tử, báo cáo hàng năm, báo cáo tài chính, các bài báo, các trang web nhằm mục đích phát hiện những thông tin sâu sắc hơn về vào hiện tượng quan tâm hoặc để củng cố thông tin từ các nguồn khác.

Các thiết kế nghiên cứu diễn giải

Nghiên cứu trường hợp. Như đã thảo luận trong chương trước, nghiên cứu trường hợp là một nghiên cứu chuyên sâu theo chiều dọc về một hiện tượng với một hoặc nhiều

địa điểm nghiên cứu với mục đích đưa ra những suy luận chi tiết, theo ngữ cảnh và tìm hiểu quá trình diễn biến phía sau hiện tượng đó. Thiết kế nghiên cứu trường hợp độc đáo ở chỗ nó có thể được sử dụng theo hướng diễn giải để xây dựng lý thuyết hoặc hướng thực chứng để kiểm tra lý thuyết. Ở chương trước, nghiên cứu trường hợp được thảo luận kỹ cả về các kỹ thuật sử dụng cũng như các ví dụ minh họa. Hơn nữa, nhà nghiên cứu trường hợp là một người quan sát trung lập (quan sát trực tiếp) chứ không phải là người tham gia tích cực (quan sát tham dự) vào sự kiện cần nghiên cứu. Cũng giống như các phương pháp diễn giải khác, quá trình hình thành nên những suy luận có ý nghĩa từ nghiên cứu trường hợp phụ thuộc rất nhiều vào các kỹ năng quan sát và khả năng tích hợp của nhà nghiên cứu.

Nghiên cứu hành động. Nghiên cứu hành động là một thiết kế nghiên cứu định tính, nhưng thực chứng nhằm mục đích kiểm nghiệm lý thuyết chứ không phải là xây dựng lý thuyết (được thảo luận trong chương này bởi vì thiếu một không gian thích hợp). Đây là một thiết kế tương tác, cho rằng rằng các hiện tượng xã hội phức tạp được nhận thức bằng cách áp dụng các thay đổi, sự can thiệp hoặc những “tác động” vào khách thể nghiên cứu và quan sát các kết quả của những hành động trên đối với khách thể nghiên cứu.

Trong phương pháp này, các nhà nghiên cứu thường là một nhà tư vấn hoặc một thành viên của tổ chức dân thân vào bối cảnh nghiên cứu (chẳng hạn như là một tổ chức). Nhà nghiên cứu khởi xướng một kế hoạch hành động để giải quyết một vấn đề và xem xét hành động đó ảnh hưởng như thế nào đến khách thể nghiên cứu, từ đó phân tích và tạo ra những hiểu biết về mối quan hệ giữa hành động và đối tượng. Ví dụ, hành động có thể bao gồm các chương trình tái cơ cấu tổ chức chẳng hạn như giới thiệu cơ cấu tổ chức, nhân sự, thủ tục hay cơ sở vật chất mới nhằm mục đích cải thiện hiệu quả làm việc của tổ chức.

Việc lựa chọn hành động của nhà nghiên cứu phải dựa trên một lý thuyết có liên quan, nên giải thích lý do tại sao và làm thế nào những hành động như vậy có thể mang lại những thay đổi như mong muốn. Giá trị thực tiễn của lý thuyết này được đánh giá bởi mức độ thành công mà hành động được lựa chọn đã khắc phục các vấn đề cần xử lý. Hiểu được vấn đề một cách sâu sắc đồng thời với việc giải quyết được vấn đề đó là tính năng trung tâm, phân biệt nghiên cứu hành động với các phương pháp nghiên cứu khác (phương pháp này hiểu được vấn đề nhưng lại không quan tâm tới việc giải quyết vấn đề) và phân biệt với hoạt động tư vấn (giải quyết vấn đề nhưng không quan tâm nhiều lý thuyết xung quanh vấn đề đó). Do đó, nghiên cứu hành động là một phương pháp tuyệt vời làm cầu nối nghiên cứu và thực tiễn.

Có vài biến thể của phương pháp nghiên cứu hành động. Phổ biến nhất là nghiên cứu hành động tham dự, được thiết kế bởi Susman và Evered (1978)¹³. Phương pháp này bao gồm một chu kỳ nghiên cứu hành động với năm giai đoạn: (1) chẩn đoán, (2) lập kế hoạch hành động, (3) thực hiện kế hoạch, (4) đánh giá, và (5) học hỏi (xem hình 10.1). *Chẩn đoán* (diagnosing) là việc nhận diện và tìm hiểu vấn đề trong bối cảnh. *Lập kế hoạch hành động* (action planning) liên quan đến việc hoạch định và đánh giá các giải pháp khác nhau cho vấn đề này và quyết định một chuỗi hành động sẽ thực hiện (dựa trên lý thuyết). *Thực hiện kế hoạch* (action taking) là việc thực hiện chuỗi hành động đã nêu trong kế hoạch. Giai đoạn *đánh giá* (evaluation) xem xét mức độ thành công của sáng kiến trong việc giải quyết vấn đề ban đầu, tức là liệu rằng giải pháp lý thuyết có thực sự hiệu quả trên thực tế. Trong giai đoạn *học hỏi* (learning), những quan sát và thông tin phản hồi trong quá trình đánh giá được kết hợp để tạo ra những hiểu biết về vấn đề, đề xuất những điều chỉnh hoặc cải tiến hành động trong tương lai.

Dựa trên kết quả đánh giá và học hỏi, hành động có thể được sửa đổi hoặc điều chỉnh để giải quyết tốt hơn các vấn đề và chu kỳ nghiên cứu hành động được lặp đi lặp lại với chuỗi hành động đã sửa đổi. Người ta khuyến nghị rằng toàn bộ chu trình nghiên cứu hành động nên được thực hiện ít nhất hai lần, có như vậy những bài học từ chu kỳ đầu tiên có thể được áp dụng trong chu kỳ thứ hai. Hình thức thu thập dữ liệu chính trong nghiên cứu hành động là quan sát tham dự, tuy vậy các kỹ thuật khác như phỏng vấn và điều tra văn bản có thể được sử dụng để củng cố thêm kết quả nghiên cứu.



Hình 12.1. Chu trình nghiên cứu hành động

¹³ Susman, G.I. and Evered, R.D. (1978). "An Assessment of the Scientific Merits of Action Research", *Administrative Science Quarterly*, (23), 582-603.

Dân tộc học (ethnography). Phương pháp nghiên cứu dân tộc học có nguồn gốc chủ yếu từ lĩnh vực nhân chủng học. Phương pháp này nhấn mạnh bối cảnh văn hóa khi nghiên cứu một hiện tượng diễn ra trong bối cảnh đó. Nhà nghiên cứu phải đắm mình trong nền văn hóa xã hội trong một thời gian dài (thường là từ 8 tháng đến 2 năm), tham gia, quan sát, ghi lại cuộc sống hàng ngày của khách thể nghiên cứu bối cảnh tự nhiên của nền văn hóa đó. Cách thức thu thập dữ liệu cơ bản trong nghiên cứu dân tộc học là quan sát tham dự và việc phân tích dữ liệu dựa trên hướng tiếp cận “tìm hiểu ý nghĩa”.

Ngoài ra, các nhà nghiên cứu phải sử dụng kỹ năng bút ký chuyên sâu và tường thuật lại những trải nghiệm của mình bằng những mô tả chi tiết để giúp độc giả có thể cảm nhận được nền văn hóa tương tự như các nhà nghiên cứu. Trong phương pháp này, nhà nghiên cứu có hai vai trò: dựa trên kiến thức chuyên môn và sự tham gia của mình vào nền văn hóa để hình thành lý thuyết và thuyết phục cộng đồng khoa học về bản chất đa dạng của hiện tượng nghiên cứu.

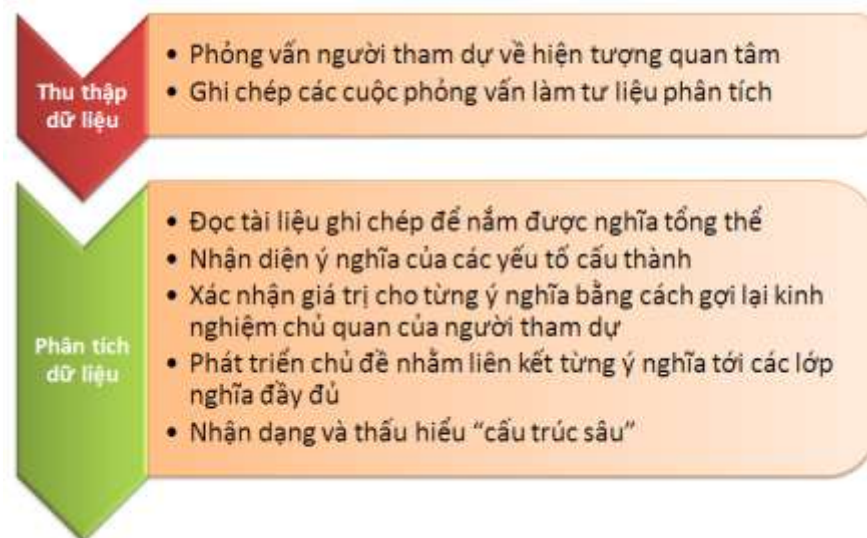
Ví dụ kinh điển hình của nghiên cứu dân tộc học là dự án nghiên cứu hành vi của động vật linh trưởng được Jane Goodall thực hiện khi cô đã sống với những con tinh tinh trong môi trường sống tự nhiên của chúng tại công viên quốc gia Gombe ở Tanzania, quan sát hành vi của chúng, giao tiếp và chia sẻ cuộc sống với chúng. Trong quá trình đó, cô đã biết được và ghi chép lại cách thức tinh tinh tìm kiếm thức ăn, trú ẩn, cách thức chúng giao tiếp, thông tin liên lạc, hành vi giao phối,...

Một ví dụ gần đây hơn là nghiên cứu dân tộc học của Myra Bluebond-Langer (1996)¹⁴ về việc ra quyết định trong những gia đình có trẻ em bị bệnh hiểm nghèo; ảnh hưởng của những vấn đề về thể chất, tinh thần, môi trường, đạo đức, pháp lý và văn hóa đến việc ra quyết định trong các gia đình đó. Trong khoảng thời gian hơn hai năm, Myra Bluebond - Langer đã tiếp xúc với khoảng 80 trẻ em mắc bệnh nan y và gia đình của các em. Để thu thập dữ liệu, nhà nghiên cứu tiến hành quan sát tham dự cuộc trò chuyện của những đứa trẻ, cha mẹ và người thân của chúng, các nhà cung cấp dịch vụ chăm sóc sức khỏe để tìm hiểu về những trải nghiệm sống của họ.

Hiện tượng học (phenomenology). Hiện tượng học là một phương pháp nghiên cứu nhấn mạnh kinh nghiệm và nhận thức của con người, là một cách để hiểu biết thực tế xung quanh chúng ta. Nó dựa trên ý tưởng của nhà triết học Đức Edmund Husserl đầu thế kỷ

¹⁴ Bluebond-Langer, M. (1996). *In the Shadow of Illness: Parents and Siblings of the Chronically Ill Child*. Princeton, NJ: Princeton University Press.

XX khi ông tin rằng kinh nghiệm của con người là nguồn gốc của tất cả các kiến thức. Hiện tượng học quan tâm tới sự phản ánh và phân tích có hệ thống các hiện tượng trong mối liên hệ với những kinh nghiệm có ý thức của con người, chẳng hạn như phán đoán, nhận thức và hành động. Nghiên cứu hiện tượng học nhằm mục đích (1) đánh giá và mô tả thực tại xã hội từ các quan điểm chủ quan khác nhau của những người có liên quan, (2) tìm hiểu ý nghĩa tượng trưng ("cấu trúc sâu") ẩn phía sau những kinh nghiệm chủ quan đó. Hiện tượng học đòi hỏi các nhà nghiên cứu phải loại bỏ bất kỳ giả định ban đầu và những thành kiến cá nhân, cảm thông với hoàn cảnh của người tham gia và nắm bắt được những khía cạnh đang thực sự tồn tại chi phối cuộc sống của họ. Nhờ vậy nhà nghiên cứu hoàn toàn có thể hiểu được cơ chế thúc đẩy, kim hãm suy nghĩ, cảm giác và hành vi có ý thức của khách thể nghiên cứu.



Hình 12.2. Phương pháp nghiên cứu hiện tượng học hiện sinh

Một số nhà nghiên cứu cho rằng hiện tượng học là một triết lý chứ không phải là một phương pháp nghiên cứu. Đáp lại bình luận này, Giorgi và Giorgi (2003)¹⁵ đã phát triển một *phương pháp nghiên cứu hiện tượng học hiện sinh* giúp định hướng cho các nghiên cứu trong lĩnh vực hiện tượng học. Như được minh họa trong hình 12.2, phương pháp này có thể được chia thành hai nhóm giai đoạn: thu thập dữ liệu và phân tích dữ liệu. Trong giai đoạn thu thập dữ liệu, những người có liên quan tới hiện tượng nghiên cứu được phỏng vấn để nắm bắt kinh nghiệm và quan điểm chủ quan của họ về hiện tượng. Ví dụ có thể hỏi “Bạn có thể mô tả một ngày thông thường của bạn được không?” hoặc “Bạn có thể

¹⁵ Giorgi, A and Giorgi, B (2003) Phenomenology. In J A Smith (ed.) Qualitative Psychology: A Practical Guide to Research Methods. London: Sage Publications.

mô tả sự kiện đó chi tiết hơn?”. Nội dung cuộc phỏng vấn được ghi chép lại cho các phân tích sau này.

Trong quá trình phân tích dữ liệu, các nhà nghiên cứu đọc văn bản chứa nội dung các cuộc phỏng vấn (biên bản phỏng vấn - transcript) để (1) có được thông tin sơ bộ về hiện tượng nghiên cứu, và (2) thiết lập “các đầu mối quan trọng” có thể mô tả một cách xác thực những kinh nghiệm chủ quan của người tham gia. Ví dụ về đầu mối quan trọng là các khái niệm như “cảm nhận về không gian” và “cảm nhận về thời gian”, sau đó những khái niệm này được dùng để thăm tra những trải nghiệm tâm lý của người tham gia. Ví dụ, liệu người tham gia cảm thấy an toàn, thoải mái, vui vẻ, hay bức xúc khi trải nghiệm một sự kiện nào đó (“cảm nhận về không gian”)? Khi tham gia vào sự kiện liệu họ có cảm thấy rằng họ bị áp lực, chậm chạp hoặc gián đoạn (“cảm nhận về thời gian”)? Phân tích hiện tượng học cần xem xét bối cảnh thời gian của người tham gia (ví dụ cảm giác của họ trong quá khứ, hiện tại và tương lai), nhà nghiên cứu phải đặt mình trong tình huống của người tham gia (tức là, tạm thời sống cuộc sống của người tham gia). Kinh nghiệm sống của người tham gia được mô tả dưới hình thức tường thuật hoặc dưới dạng các chủ đề quan trọng xuất hiện sau quá trình phân tích. Sau đó các chủ đề được khai thác sâu để xác định các tầng nghĩa khác nhau từ thông tin của người tham gia, đồng thời vẫn phải giữ lại được sự tinh tế và tính đa sắc thái của thông tin đó.

Tính xác thực và giá trị khoa học trong nghiên cứu diễn giải

Nghiên cứu thực chứng sử dụng hướng tiếp cận "giản lược" (reductionist) thông qua quá trình đơn giản hoá các thực tại xã hội phức tạp thành các lý thuyết ngắn gọn, trong khi đó nghiên cứu diễn giải lại nỗ lực giải thích các thực tại xã hội thông qua quan điểm chủ quan của những người tham gia sau khi họ đã có những trải nghiệm sâu sắc trong bối cảnh thực tế. Do vậy, việc giải thích này sẽ bị ràng buộc chặt chẽ bởi bối cảnh nghiên cứu và vì vậy chúng rất khó để có thể khái quát hóa trong những bối cảnh khác. Đồng thời, do phân tích diễn giải mang ý chí chủ quan và phụ thuộc vào hiểu biết và kinh nghiệm của người nghiên cứu nên các nhà thực chứng (hay chức năng luận) thường chỉ trích phân tích diễn giải là thiếu khách quan, thiếu chặt chẽ.

Bởi vì nghiên cứu diễn giải dựa trên các giả định bản thể luận và nhận thức luận về hiện tượng xã hội khác xa về bản chất so với nghiên cứu thực chứng, các đánh giá về tính khách quan, chặt chẽ, giá trị khoa học ở nghiên cứu thực chứng, chẳng hạn như độ tin cậy, giá trị nhân quả và tính phổ quát, không áp dụng một cách tương tự như với nghiên cứu

diễn giải. Tuy nhiên, Lincoln và Guba (1985)¹⁶ đã đưa ra một bộ tiêu chuẩn thay thế có thể được sử dụng để đánh giá giá trị khoa học của nghiên cứu diễn giải.

Tính ổn định (dependability). Nghiên cứu diễn giải có thể được xem như là có tính ổn định nếu hai nhà nghiên cứu đánh giá cùng một hiện tượng, sử dụng cùng một hệ thống luận chứng, một cách độc lập, họ đưa ra những kết luận tương tự; hoặc một nhà nghiên cứu quan sát cùng một hiện tượng hoặc một hiện tượng tương tự ở những thời điểm khác nhau, nhưng mang đến những kết luận tương tự. Tiêu chí này tương tự như độ tin cậy (reliability) trong nghiên cứu thực chứng nếu như hai nhà nghiên cứu độc lập đưa ra khái niệm tương tự về *độ tin cậy nội bộ* (inter-rater reliability) hoặc nếu cùng một nhà nghiên cứu thực hiện hai quan sát tương tự về một hiện tượng và đạt được *độ tin cậy liên tục* (test-retest reliability). Để đảm bảo tính ổn định, các nhà nghiên cứu diễn giải phải cung cấp chi tiết, đầy đủ về hiện tượng quan tâm và bối cảnh xã hội mà nó gắn bó, điều này cho phép độc giả xác thực một cách độc lập những suy luận diễn giải của nhà nghiên cứu.

Tín nhiệm (credibility). Nghiên cứu diễn giải có thể được tin nhiệm nếu người đọc cảm thấy tin vào các kết luận của nghiên cứu. Khái niệm về tín nhiệm này giống như khái niệm *giá trị nhân quả* (internal validity) trong nghiên cứu chức năng luận. Độ tín nhiệm của nghiên cứu diễn giải có thể được cải thiện bằng cách cung cấp bằng chứng về tham gia sâu rộng của nhà nghiên cứu trong lĩnh vực đang nghiên cứu, bằng cách thể hiện kỹ năng phối kết hợp các dữ liệu đa dạng từ nhiều nguồn với nhiều kỹ thuật thu thập và phân tích, bằng cách thực thi việc quản lý và phân tích dữ liệu tỉ mỉ, như sao chép nguyên văn nội dung cuộc phỏng vấn sang biên bản phỏng vấn, lưu giữ cẩn thận, chính xác các địa chỉ liên lạc và các cuộc phỏng vấn, ghi chú rõ ràng các lựa chọn về lý thuyết và phương pháp luận. Tất cả những công việc này cho phép một cuộc kiểm định độc lập về quá trình thu thập và phân tích dữ liệu nếu cần thiết.

Xác thực (confirmability). Tính xác thực đề cập đến mức độ mà các kết quả báo cáo trong nghiên cứu diễn giải có thể được xác thực một cách độc lập bởi những người khác (thông thường là những người tham gia). Tiêu chí này tương tự như khái niệm về *tính khách quan* (objectivity) trong nghiên cứu chức năng luận. Bởi vì nghiên cứu diễn giải bác bỏ ý niệm về một thực thể khách quan, nên tính xác thực của nó được thể hiện ở góc độ “liên chủ quan” (inter-subjectivity), ví dụ nếu những người tham gia nghiên cứu đồng tình

¹⁶ Lincoln, Y. S., and Guba, E. G. (1985). *Naturalistic Inquiry*. Beverly Hills, CA: Sage Publications.

với những kết luận của nhà nghiên cứu (dựa trên việc xem xét các báo cáo của đề tài nghiên cứu), thì phát hiện trong nghiên cứu có thể được xem là xác thực.

Khái quát (transferability). Khả năng khái quát trong nghiên cứu diễn giải đề cập đến mức độ mà các kết quả có thể được khái quát hóa cho các đối tượng khác. Tiêu chí này tương tự như *giá trị mở rộng* (*external validity*) trong các nghiên cứu chức năng luận. Nhà nghiên cứu phải cung cấp những mô tả phong phú, chi tiết về bối cảnh nghiên cứu ("mô tả sâu"), giải thích thấu đáo, toàn diện về các phạm trù, giả thiết và lộ trình nghiên cứu để độc giả có thể đánh giá một cách độc lập về khả năng và mức độ áp dụng các kết quả nghiên cứu cho các đối tượng khác hay một bối cảnh khác.

Chương 13

PHÂN TÍCH ĐỊNH TÍNH

Phân tích định tính là việc phân tích những dữ liệu định tính như dữ liệu bằng văn bản được sao chép lại từ những cuộc phỏng vấn (biên bản phỏng vấn). Không giống như phân tích định lượng, đó là việc thống kê số liệu và phần lớn là độc lập với nhà nghiên cứu, phân tích định tính phụ thuộc nhiều vào kỹ năng phân tích tổng hợp, kiến thức chuyên môn và hiểu biết xã hội về khách thể nghiên cứu của cá nhân nhà nghiên cứu. Phân tích định tính nhấn mạnh nhiệm vụ “ý nghĩa hóa” hoặc “tri thức hóa” một hiện tượng, chứ không phải là dự đoán hay giải thích nó. Tư duy sáng tạo, quyết tâm khám phá và tinh thần cởi mở là chìa khóa mang đến những phân tích định tính có hàm lượng khoa học cao. Chương này sẽ giới thiệu tổng quan về một số phương pháp phân tích định tính. Bạn đọc quan tâm đến nội dung này có thể tìm đọc những cuốn sách chuyên khảo ví dụ như sản phẩm của Miles và Huberman (1984).¹⁷

Lý thuyết nền

Làm thế nào bạn có thể phân tích một tập hợp dữ liệu định tính không lồ thu được thông qua các quan sát tham dự, các cuộc phỏng vấn chuyên sâu, hội thảo nhóm nhỏ, những câu chuyện từ bản ghi âm, video hoặc các tài liệu thứ cấp? Một trong những kỹ thuật để phân tích các dữ liệu văn bản này là **lý thuyết nền**^(*) (grounded theory) - một kỹ thuật quy nạp giúp giải nghĩa những dữ liệu đã lưu giữ về một hiện tượng xã hội để xây dựng lý thuyết về hiện tượng đó. Kỹ thuật này được Glaser và Strauss (1967)¹⁸ phát triển khi giới thiệu *phương pháp phân tích so sánh liên tiếp* (method of constant comparative analysis) trong nghiên cứu về thiết lập lý thuyết. Phương pháp này tiếp tục được hoàn thiện bởi Strauss và Corbin (1990)¹⁹ khi họ đề xuất các kỹ thuật cụ thể về *mã hóa* (*coding* - một quá trình phân loại các nhóm dữ liệu văn bản thành các nhóm ngữ nghĩa (các khái niệm), các phạm trù và các mối quan hệ. Trong quá trình nghiên cứu, lý thuyết sẽ được xây dựng

¹⁷ Miles M. B., Huberman A. M. (1984). *Qualitative Data Analysis: A Sourcebook of New Methods*. Newbury Park, CA: Sage Publications.

^(*) Grounded theory có khi được dịch sang tiếng Việt là: *lý thuyết tiềm ẩn; lý thuyết xây dựng trên nền tảng dữ liệu* (N.D).

¹⁸ Glaser, B. and Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*, Chicago: Aldine.

¹⁹ Strauss, A. and Corbin, J. (1990). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, Beverly Hills, CA: Sage Publications.

dựa trên cơ sở dữ liệu quan sát thực nghiệm, đó là lý do đặt tên cho kỹ thuật này là lý thuyết nền (hay lý thuyết xây dựng trên nền tảng dữ liệu). Để đảm bảo lý thuyết chỉ hình thành dựa trên những thông tin đã quan sát, *lý thuyết nền* đòi hỏi các nhà nghiên cứu loại phải bỏ tất cả các phỏng đoán ban đầu hoặc những thiên vị cá nhân trước khi bắt đầu phân tích dữ liệu, hãy để cho dữ liệu quyết định việc hình thành lý thuyết.

Strauss và Corbin (1998) mô tả ba kỹ thuật mã hoá để phân tích dữ liệu văn bản: mã hoá mở, mã hoá trục và mã hoá chọn lọc. **Mã hoá mở** (*open coding*) là một quá trình giúp nhận diện các khái niệm hoặc các ý tưởng quan trọng ẩn bên trong dữ liệu văn bản mà chúng có khả năng liên quan đến hiện tượng đang nghiên cứu. Nhà nghiên cứu đọc các dữ liệu sống theo từng dòng để phát hiện các sự kiện, tình tiết, ý tưởng, hành động, nhận thức riêng lẻ rời rạc và các mối tương tác có liên quan; chúng sẽ định hình thành các khái niệm (vì thế được gọi là các mã hoá sống). Từng khái niệm được liên kết tới các phần cụ thể của văn bản, gọi các đơn vị mã hoá (coding unit) phục vụ cho việc củng cố sau này.

Một số khái niệm có thể được hình thành một cách đơn giản và rõ ràng, trong khi những khái niệm khác có thể phức tạp, mơ hồ và được nhìn nhận khác nhau bởi những người tham gia khác nhau. Các đơn vị mã hoá có thay đổi với các khái niệm được trích chép. Các khái niệm đơn giản như "quy mô tổ chức" có thể chỉ bao gồm một vài từ trong văn bản, trong khi những khái niệm phức tạp như "sứ mệnh của tổ chức" có thể kéo dài nhiều trang. Khái niệm có thể được đặt tên bằng cách sử dụng quy ước đặt tên của riêng nhà nghiên cứu hoặc các dán nhãn được tiêu chuẩn hóa từ các nghiên cứu trước. Khi một tập hợp các khái niệm cơ bản được xác định, các khái niệm này sau đó có thể được sử dụng để mã hoá phần còn lại của dữ liệu, trong khi đó vẫn không bỏ qua việc tìm kiếm những khái niệm mới và tinh chỉnh các khái niệm đã hình thành. Trong khi định ngữ, việc xác định các đặc điểm dễ nhận biết, đặc thù của từng khái niệm là điều hết sức quan trọng, chẳng hạn như kích thước, màu sắc hoặc mức độ (ví dụ như, cao hay thấp) của khái niệm đó, có như vậy sau này các khái niệm tương tự có thể được nhóm lại với nhau. Kỹ thuật mã hoá này được gọi là "mở" bởi vì các nhà nghiên cứu mở ra và tích cực tìm kiếm các khái niệm mới có liên quan đến hiện tượng quan tâm.

Tiếp theo, các khái niệm tương tự được nhóm lại thành mức cao hơn gọi là các *phân nhóm* (category). Trong khi khái niệm sẽ được tìm thấy trong bối cảnh cụ thể, phân nhóm có xu hướng mở rộng và khái quát hơn và cuối cùng phát triển thành các *phạm trù* (construct) trong kỹ thuật thiết lập lý thuyết. Để hình thành các phân nhóm, việc giảm số lượng các khái niệm là cần thiết. Muốn vậy các nhà nghiên cứu phải xác định những vấn

đề cốt yếu về hiện tượng đang nghiên cứu bằng cách phác họa một "bức tranh tổng thể" nêu bật những vấn đề đó.

Phân nhóm có thể được hình thành qua nhiều giai đoạn, ví dụ, đầu tiên kết hợp các khái niệm thành các cụm và sau đó kết hợp các cụm thành các phân nhóm. Các phạm trù được đề cập trong các nghiên cứu trước đó có thể được sử dụng để đặt tên cho các phân nhóm này, đặc biệt nếu mục tiêu của nghiên cứu là để mở rộng lý thuyết hiện có. Tuy nhiên, phải rất thận trọng trong khi sử dụng các phạm trù trước đó, bởi vì chúng có thể đi kèm với những quan niệm truyền thống và thành kiến phổ biến.

Đối với mỗi phân nhóm, cần phân biệt các đặc trưng (characteristics or properties) của phân nhóm với các khía cạnh (dimensions) của phân nhóm đó. Khía cạnh đề cập đến giá trị của một đặc trưng tồn tại trong thời gian dài và liên tục. Ví dụ, phân nhóm "phương tiện thông tin liên lạc" có thể có một đặc trưng là "tốc độ" và đặc trưng "tốc độ" này có thể có một số khía cạnh như nhanh, trung bình hay chậm. Việc phân loại như vậy giúp phân biệt giữa các loại hình khác nhau của các phương tiện thông tin liên lạc và cho phép nhà nghiên cứu xác định các *mô hình* (patterns) trong dữ liệu, chẳng hạn như phương tiện truyền thông nào được sử dụng cho loại nhiệm vụ nào.

Giai đoạn thứ hai của thiết lập lý thuyết là **mã hoá theo trục** (axial coding), trong đó các phân nhóm và cụm được kết hợp tạo thành các mối quan hệ nhân quả hoặc các giả thuyết có thể giải thích các hiện tượng quan tâm. Mặc dù khác với mã hoá mở, nhưng mã hoá theo trục có thể thực hiện đồng thời với mã hoá mở. Mối quan hệ giữa các phân nhóm có thể nhìn thấy rõ từ trong dữ liệu hoặc cũng có thể ẩn và tinh tế. Trong trường hợp các mối quan hệ là không rõ ràng, các nhà nghiên cứu có thể sử dụng lược đồ mã hoá (coding scheme, thường được gọi là "mô thức mã hoá (coding paradigm)", nhưng khác với các mô thức được thảo luận trong Chương 3) để xác định đâu là *điều kiện* (conditions - tình huống chứa đựng hiện tượng nghiên cứu), đâu là *hành động/ tương tác* (actions/ interactions - sự phản ứng, tương tác của các cá nhân đối với các sự kiện trong điều kiện đã xác định) và đâu là *kết quả* (consequences - các kết quả của hành động). Khi các điều kiện, tương tác và kết quả được xác định, các luận điểm lý thuyết bắt đầu xuất hiện và nhà nghiên cứu có thể bắt đầu giải thích lý do tại sao một hiện tượng xảy ra, trong những điều kiện nào, gây ra những gì hậu quả gì.

Giai đoạn thứ ba và cuối cùng của thiết lập lý thuyết là *mã hoá chọn lọc* (selective coding) nhằm xác định một phân nhóm trung tâm (hay một biến cốt lõi) và sau đó liên kết

chúng một cách có hệ thống và logic với các phân nhóm khác. Phân nhóm trung tâm (central category) có thể được phát triển từ việc kết hợp các phân nhóm hiện có. Dữ liệu mới được cần được thu thập có chọn lọc để củng cố phân nhóm trung tâm và các mối quan hệ của nó với các phân nhóm khác (đây có thể là một lý thuyết dự kiến).

Mã hoá chọn lọc giúp thu nhỏ phạm vi phân tích và đẩy nhanh quá trình phân tích. Đồng thời, nhà nghiên cứu phải chú ý tới sự xuất hiện của các phạm trù mới từ các dữ liệu mới liên quan đến hiện tượng quan tâm (liên quan đến mã hoá mở), điều này có thể giúp phát triển hơn nữa lý thuyết dự kiến ban đầu. Do đó, mã hoá mở, mã hoá theo trục và mã hoá chọn lọc có thể được tiến hành đồng thời. Việc mã hoá các dữ liệu mới và phát triển lý thuyết vẫn phải tiếp tục cho đến khi đạt đến độ *bão hòa lý thuyết* (*theoretical saturation*), nghĩa là, khi các dữ liệu bổ sung không mang lại bất kỳ sự thay đổi đáng kể nào trong các phân nhóm trung tâm và mối quan hệ của chúng.

Quá trình "so sánh liên tục" đề cập đến việc sắp xếp lại, tập hợp và sàng lọc liên tiếp các phân nhóm, mối quan hệ và các giả thuyết. Quá trình này dựa vào mức độ hiểu biết ngày một sâu sắc và mối quan hệ tương hỗ của bốn hành vi: (1) so sánh các tình tiết / văn bản liên quan tới mỗi phạm trù (để củng cố phạm trù đó), (2) hợp nhất các phạm trù và các thuộc tính của chúng, (3) khoanh vùng, chọn lọc lý thuyết (tập trung vào các khái niệm cốt lõi và bỏ qua khái niệm ít liên quan hơn), và (4) hình thành lý thuyết (bằng cách sử dụng các kỹ thuật như bình luận, tường thuật và đồ thị hóa được thảo luận trong chương kế tiếp).

Việc có một phân nhóm trung tâm không có nghĩa là tất cả các phân nhóm khác có thể được liên kết dễ dàng xung quanh phân nhóm này. Để xác định đâu là điều kiện, tương tác hay kết quả của phân nhóm trung tâm, Strauss và Corbin (1990) khuyến nghị sử dụng một số kỹ thuật tích hợp như tường thuật, bình luận và đồ thị hóa. Trong *kỹ thuật tường thuật* (storylining), các phân nhóm và các mối quan hệ được sử dụng để giải thích và điều chỉnh một tường thuật về hiện tượng quan sát. Bình luận là các bài viết lý thuyết có chứa các khái niệm cơ bản cũng như các mối quan hệ được xác định trong quá trình phân tích lý thuyết nên, nó là công cụ quan trọng để phát hiện và điều chỉnh các ý tưởng trong quá trình phân tích. *Kỹ thuật bình luận* (memoing) trong thiết lập lý thuyết được sử dụng để khám phá mô hình các mối quan hệ giữa các phân nhóm nhờ sử dụng từng cặp bảng, cặp sơ đồ, cặp con số, hay các hiển thị minh họa khác. *Kỹ thuật sơ đồ hóa khái niệm* (concept mapping) là việc diễn đạt bằng sơ đồ các khái niệm và các mối quan hệ giữa chúng (ví dụ, sử dụng các hộp và các mũi tên). Các khái niệm cơ bản thường được bố trí trên một hoặc

nhiều tờ giấy, bảng đen hoặc sử dụng chương trình phần mềm đồ họa, liên kết với nhau bằng các mũi tên, được điều chỉnh sao cho phù hợp nhất với dữ liệu đã thu thập.

Phân tích nội dung

Phân tích nội dung (content analysis) là phân tích có hệ thống các nội dung trong văn bản (ví dụ, người đó nói gì, cho ai, tại sao, đến mức độ nào và ảnh hưởng của nó) theo hướng định lượng hoặc định tính. Phân tích nội dung thường được thực hiện như sau. Đầu tiên, cần tìm kiếm các bản văn để phân tích (ví dụ như những câu chuyện trên báo, báo cáo, bài viết trên blog, nhận định trực tuyến, v.v...), nhà nghiên cứu bắt đầu bằng việc chọn một mẫu gồm tập hợp các văn bản có chọn lọc. Đây này không phải là quá trình chọn ngẫu nhiên mà nhà nghiên cứu phải lựa chọn một cách có chọn lọc các văn bản có nội dung phù hợp.

Thứ hai, nhà nghiên cứu xác định và áp dụng các quy tắc để phân chia văn bản thành các mảng hoặc các "khối", chúng có thể được coi như các đơn vị phân tích riêng biệt. Quá trình này được gọi là giai đoạn *hình thành* (unitizing) đơn vị phân tích. Ví dụ, từ các văn bản có thể tạo thành các nhóm như nhóm giả định, nhóm hậu quả, nhóm tác động và nhóm rào cản. Thứ ba, nhà nghiên cứu xây dựng và áp dụng một hoặc nhiều khái niệm từ mỗi đoạn văn bản đã phân nhóm theo một quá trình gọi là mã hoá. Để mã hoá, nhà nghiên cứu sử dụng một lược đồ mã hoá được thiết kế dựa trên các chủ điểm đang tìm kiếm. Cuối cùng, các dữ liệu đã mã hoá được phân tích, thường theo cả hướng định tính và định lượng, nhằm xác định chủ điểm nào thường xuất hiện nhất, trong điều kiện nào và chúng liên kết với nhau như thế nào.

Một hình thức đơn giản của phân tích nội dung là ***phân tích cảm tính*** (sentiment analysis) - kỹ năng này được sử dụng để nắm bắt quan điểm hoặc thái độ của người dân đối với một đối tượng, một người hoặc hiện tượng. Đọc tin nhắn trực tuyến về một ứng cử viên chính trị được đăng trên một diễn đàn và phân loại mỗi tin nhắn là tích cực, tiêu cực hoặc trung lập là một ví dụ về phân tích này. Trong trường hợp này, mỗi tin nhắn đại diện cho một đơn vị phân tích. Phân tích như vậy sẽ giúp nhận định liệu các cử tri tỏ thái độ tích cực, tiêu cực hay trung lập đối với ứng cử viên đó. Những phân tích như thế này có thể được thực hiện một cách thủ công bằng tay. Tuy nhiên, đối với những nguồn dữ liệu rất lớn (hàng triệu hồ sơ văn bản) và nguồn biến động theo thời gian, các chương trình phần mềm xử lý ngôn ngữ và phân tích văn bản có thể được sử dụng để tự động hóa quá trình mã hoá và lưu giữ hồ sơ.

Yếu điểm bị chỉ trích thường xuyên trong phân tích nội dung đó là nó thiếu một quy trình có hệ thống giúp phân tích này có thể được áp dụng rộng rãi. Schilling (2006)²⁰ khắc phục chỉ trích này bằng cách tổ chức các bước phân tích nội dung thành một mô hình xoắn ốc. Mô hình này bao gồm năm giai đoạn: (1) chuyển đổi các băng ghi âm thành dữ liệu văn bản thô, (2) chuyển đổi dữ liệu thô thành bản thảo cô đọng, (3) chuyển đổi bản thảo cô đọng thành các phân nhóm sơ bộ, (4) sử dụng các phân nhóm sơ bộ để tạo ra các giao thức mã hoá, và (5) phân tích các giao thức mã hoá để đưa ra các giải thích về hiện tượng quan tâm.

Phân tích nội dung có một vài hạn chế. Đầu tiên, quá trình mã hoá bị giới hạn chỉ trong những thông tin có sẵn ở dạng văn bản. Ví dụ, nếu một nhà nghiên cứu quan tâm đến việc nghiên cứu quan điểm của người dân về án tử hình, nhưng không có hồ sơ lưu trữ các tài liệu văn bản có sẵn, vậy thì phân tích không thể được thực hiện. Thứ hai chọn mẫu phải được thực hiện một cách cẩn thận để tránh thiên vị. Ví dụ, nếu khách thể nghiên cứu của bạn là tài liệu nghiên cứu đã công bố về một chủ đề nhất định, thì bạn đã bỏ qua toàn bộ hệ thống những nghiên cứu không công bố hoặc đề tài gần đây chưa được công bố.

Phân tích thông diễn

Phân tích thông diễn (hermeneutic analysis) là một loại đặc biệt của phân tích nội dung, trong đó nhà nghiên cứu cố gắng "diễn đạt" ý nghĩa chủ quan của một văn bản trong bối cảnh xã hội lịch sử của nó. Không giống như các kỹ thuật thiết lập lý thuyết hoặc hay phân tích nội dung chúng bỏ qua bối cảnh và ý nghĩa của các tài liệu văn bản trong quá trình định ngữ, phân tích thông diễn là một kỹ thuật đậm chất diễn giải khi phân tích dữ liệu định tính. Phương pháp này giả định rằng các văn bản thuật lại trải nghiệm của tác giả trong một bối cảnh lịch sử - xã hội nhất định nên được diễn giải trong bối cảnh đó. Vì vậy, nhà nghiên cứu cần liên tục đưa ra và liên kết những giải thích riêng lẻ từ tài liệu văn bản (một phần) và sự hiểu biết toàn diện về bối cảnh (toàn bộ) để phát triển những hiểu biết đầy đủ hơn về hiện tượng trong bối cảnh của nó. Quá trình này được triết gia người Đức Martin Heidegger gọi là *chu kỳ thông diễn*. Thuật ngữ thông diễn (ở số ít) đề cập đến một phương pháp hay một hướng giải thích cụ thể nào đó.

Nói một cách tổng quát hơn, thông diễn học là nghiên cứu về suy diễn, lý thuyết và thực hành suy diễn. Xuất phát từ những nghiên cứu tôn giáo và ngôn ngữ học, thông diễn học truyền thống, chẳng hạn *thông diễn học kinh thánh*, đề cập đến việc giải thích các văn

²⁰ Schilling, J. (2006). "On the Pragmatics of Qualitative Assessment: Designing the Process for Content Analysis," *European Journal of Psychological Assessment* (22:1), 28-37.

bản, đặc biệt là trong các lĩnh vực tôn giáo, văn học và pháp luật (như Kinh Thánh). Trong thế kỷ XX, Heidegger cho rằng, phương pháp trực tiếp hơn, không qua trung gian và xác thực hơn để chiếm lĩnh tri thức về thực tại xã hội là hãy trải nghiệm nó, chứ không phải chỉ đơn giản là quan sát nó. Vì vậy, ông đề xuất *thông diễn học triết lý* chuyển trọng tâm từ sự giải thích sang tri thức hiện sinh. Heidegger cho rằng văn bản là phương tiện mà nhờ đó người đọc không chỉ đọc về trải nghiệm của tác giả, mà còn hồi tưởng lại những trải nghiệm đó. Thông diễn học đương đại, được phát triển bởi các học trò của Heidegger như Hans-Georg Gadamer, nghiên cứu sâu hơn những hạn chế của văn bản viết trong diễn đạt trải nghiệm xã hội, đã đi đến đề xuất một nền tảng cho tiến trình diễn giải, bao gồm tất cả các hình thức giao tiếp như văn bản, lời nói và không bằng lời nói. Họ cũng đã tìm ra những vấn đề làm hạn chế khả năng giao tiếp của văn bản viết, chẳng hạn như những giả định, các cấu trúc ngôn ngữ (ví dụ như ngữ pháp, cú pháp,...) và ngữ nghĩa học (nghiên cứu các dấu hiệu trong văn bản như biểu tượng hóa, phép loại suy, lối ẩn dụ và ngôn ngữ châm biếm). Dù không hoàn toàn chính xác, nhưng thuật ngữ thông diễn học đôi khi được sử dụng thay thế cho *giải thích học* (exegesis), trong đó đề cập đến những luận giải hoặc những giải thích phê phán về các văn bản viết đặc biệt là các văn bản tôn giáo.

Kết luận

Cuối cùng, các chương trình phần mềm tiêu chuẩn, như ATLAS.ti.5, NVivo và QDA Miner, có thể được sử dụng để tự động hóa quá trình mã hoá trong các phương pháp phân tích định tính. Các chương trình này có thể tổ chức, tìm kiếm, sắp xếp và xử lý một cách nhanh chóng và hiệu quả những khối lượng lớn dữ liệu văn bản với các quy tắc lập trình do người dùng thiết lập. Để định hướng cho những phân tích tự động như vậy cần tạo ra một lược đồ mã hoá, quy định cụ thể cho các từ khóa hoặc ngữ nghĩa để tìm kiếm trong văn bản, dựa trên việc kiểm tra thủ công ban đầu một mẫu dữ liệu văn bản. Lược đồ này có thể được tổ bố trí theo cấp độ khái quát tăng dần của các khái niệm. Các lược đồ mã hoá nên được củng cố bằng cách sử dụng một mẫu văn bản khác để kiểm định tính chính xác và đầy đủ. Nếu các lược đồ mã hoá là thiên vị hoặc không chính xác thì kết quả phân tích của toàn bộ khách thể nghiên cứu có thể thiếu sót và không thể diễn giải được. Tuy nhiên, các chương trình phần mềm không thể phát hiện được tính biểu cảm của một số từ hoặc cụm từ và cũng không thể phát hiện ra ý nghĩa của bối cảnh mà văn bản được sử dụng (chẳng hạn như những cụm từ trong ngôn ngữ châm biếm hoặc lối ẩn dụ). Do vậy, các phần mềm có thể dẫn đến những suy diễn hoặc giải thích không trọn nghĩa, thậm chí sai lệch, nhất là trong phân tích định tính quy mô lớn.

Chương 14

PHÂN TÍCH ĐỊNH LƯỢNG: THỐNG KÊ MÔ TẢ

Dữ liệu số được thu thập trong một dự án nghiên cứu có thể được phân tích định lượng bằng cách sử dụng các công cụ thống kê theo hai cách khác nhau. **Phân tích mô tả** (descriptive analysis) là quá trình sử dụng các kỹ năng thống kê để mô tả, tổng hợp và trình bày các phạm trù và sự liên kết giữa chúng. **Phân tích suy luận** (inferential analysis) liên quan đến việc thử nghiệm các giả thuyết thống kê (kiểm nghiệm lý thuyết). Trong chương này, chúng ta sẽ tìm hiểu các kỹ thuật thống kê được sử dụng trong phân tích mô tả và chương tiếp theo sẽ đề cập đến phân tích suy luận. Nhiều phân tích dữ liệu định lượng hiện nay được thực hiện bằng cách sử dụng các chương trình phần mềm như SPSS (*Statistical Product and Service Solutions*)^(*) hoặc SAS (*Statistical Analysis System*)^(**). Độc giả nên làm quen với một trong những phần mềm đó để có thể hiểu được các khái niệm trong chương này.

Chuẩn bị dữ liệu

Trong các dự án nghiên cứu, dữ liệu có thể được thu thập từ nhiều nguồn khác nhau: các cuộc khảo sát qua thư, phỏng vấn, dữ liệu tiền và hậu thử nghiệm, dữ liệu quan sát,... Các dữ liệu này phải được chuyển đổi thành định dạng số để máy tính có thể đọc được, chẳng hạn như trong một bảng tính hoặc một tập tin văn bản, để từ đó chúng có thể được phân tích bằng chương trình máy tính như SPSS hoặc SAS. Chuẩn bị dữ liệu thường theo các bước sau.

Mã hóa dữ liệu (data coding). Mã hóa là quá trình chuyển đổi dữ liệu sang định dạng số. Bảng mã hóa được tạo ra để hướng dẫn quá trình mã hóa. **Bảng mã hóa** (codebook) là một tài liệu tổng hợp mô tả chi tiết từng biến số trong một nghiên cứu, các chỉ số hoặc các thước đo cho biến số, định dạng của từng mục (dạng số, văn bản, v.v.), thang trả lời cho mỗi chỉ số (tức là, nó được đo bằng thang định danh, thang thứ tự, thang khoảng hay thang tỷ lệ; thang đó là năm điểm, bảy điểm hay là thang khác) và cách thức mã hóa mỗi giá trị chỉ số theo một định dạng số. Ví dụ, nếu chúng ta đo lường một chỉ số

(*) Phần mềm SPSS là viết tắt của tên ban đầu *Statistical Package for the Social Sciences*. Tên gọi phiên bản 2015 là *IBM SPSS Statistics* (xem thêm tại địa chỉ: <http://www-01.ibm.com/software/analytics/spss/>).

(**) Phần mềm SAS là một hệ thống tích hợp các sản phẩm phần mềm của SAS Institute Inc. phát triển (xem thêm tại địa chỉ: http://www.sas.com/en_us/home.html).

theo thang Likert điểm 7 với các mấu xếp từ “hoàn toàn không đồng ý” đến “hoàn toàn đồng ý”, chúng ta có thể mã hóa chỉ số đó là: 1 = hoàn toàn không đồng ý, 4 = trung lập và 7 = hoàn toàn đồng ý và các mấu giá trị trung gian ở giữa. Dữ liệu định danh chẳng hạn như các ngành công nghiệp có thể được mã hóa ở dạng số bằng cách sử dụng một chương trình mã hóa, chẳng hạn như: 1 = sản xuất, 2 = bán lẻ, 3 = tài chính, 4 = chăm sóc sức khỏe,... (tất nhiên, dữ liệu định danh không thể phân tích thống kê được). Thang tỷ lệ như tuổi tác, thu nhập hoặc điểm thi có thể được mã hóa giống như thông tin của người trả lời đã cung cấp.

Đôi khi, dữ liệu cần được tập hợp thành các dạng khác với dạng ban đầu khi thu thập dữ liệu. Ví dụ, để đo lường một phạm trù như "lợi ích của máy tính", người trả lời được cung cấp một danh sách bao gồm các lợi ích mà họ có thể lựa chọn (tức là, họ có thể lựa chọn không giới hạn trong số những lợi ích đó theo như họ muốn), sau đó tổng số các mục được chọn có thể được sử dụng như một thước đo tổng hợp của các lợi ích. Lưu ý rằng nhiều hình thức khác của dữ liệu, chẳng hạn như biên bản phỏng vấn, không thể chuyển đổi thành định dạng số để phân tích thống kê. Công việc mã hóa là đặc biệt quan trọng đối với các nghiên cứu rộng, phức tạp bao gồm nhiều biến và nhiều chỉ số đo lường, ở đó quá trình mã hóa có thể được thực hiện bởi nhiều người khác nhau. Lúc này nhóm làm việc cần phải mã hóa, hiểu và giải thích dữ liệu một cách thống nhất.

Nhập dữ liệu (data entry). Dữ liệu mã hoá có thể được nhập vào một bảng tính, cơ sở dữ liệu, tập tin văn bản hoặc trực tiếp vào chương trình thống kê như phần mềm SPSS. Hầu hết các chương trình thống kê cung cấp một chương trình soạn thảo để nhập dữ liệu. Tuy nhiên, các chương trình này lưu trữ dữ liệu ở định dạng gốc của nó (ví dụ, SPSS lưu trữ dữ liệu bằng các tập tin dạng .sav), điều này gây khó khăn cho việc chia sẻ dữ liệu với các chương trình thống kê khác. Do đó, tốt hơn là nên nhập dữ liệu vào một bảng tính hoặc cơ sở dữ liệu, nơi chúng có thể được sắp xếp lại khi cần thiết hoặc khi muốn chia sẻ giữa các chương trình phần mềm và như vậy các tập con của dữ liệu có thể được trích xuất để phân tích.

Bộ dữ liệu nhỏ với ít hơn 65.000 quan sát và 256 chỉ số có thể được lưu trữ trong một bảng tính như Microsoft Excel, trong khi tập dữ liệu lớn hơn với hàng triệu quan sát sẽ đòi hỏi một cơ sở dữ liệu. Mỗi quan sát có thể được nhập vào một hàng trong bảng tính và từng chỉ số đo lường có thể được biểu diễn như là một cột. Phải thường xuyên kiểm tra độ chính xác của dữ liệu nhập vào bằng cách kiểm tra ngẫu nhiên các cặp chỉ số và quan sát trong và sau khi nhập. Thêm vào đó, trong khi nhập dữ liệu, người nhập nên cảnh giác khi

xuất hiện một bảng chứng rõ ràng về dữ liệu xấu, chẳng hạn như người trả lời lựa chọn "hoàn toàn đồng ý" trong tất cả các mục không phân biệt nội dung, kể cả những mục mã hóa ngược. Nếu vậy, các dữ liệu đó có thể vẫn được nhập nhưng phải loại trừ chúng khỏi phân tích tiếp theo.

Dữ liệu khuyết (missing value). Dữ liệu bị thiếu là một phần không thể tránh khỏi của bất kỳ gói dữ liệu thực nghiệm nào. Người được hỏi không thể trả lời các câu hỏi quá khó hoặc chúng được diễn đạt mơ hồ hay quá nhạy cảm. Vấn đề đó phải được phát hiện trước qua quá trình thử nghiệm và điều chỉnh trước khi chính thức bắt đầu thu thập dữ liệu. Trong quá trình nhập dữ liệu, một số chương trình thống kê sẽ tự động mặc định các mục trống nghĩa là giá trị thiếu, trong khi đó, những phần mềm khác đòi hỏi một giá trị cụ thể chẳng hạn là -1 hoặc 999 phải được nhập vào để biểu thị cho một giá trị bị thiếu.

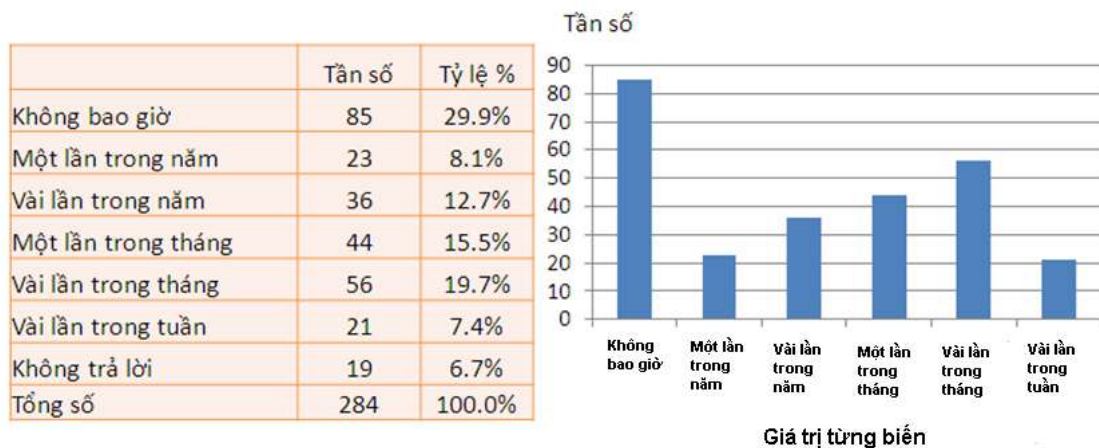
Trong quá trình phân tích dữ liệu, chế độ mặc định xử lý các giá trị bị thiếu trong hầu hết các chương trình phần mềm chỉ đơn giản là đưa toàn bộ các quan sát có chứa từ giá trị thiếu vào trong một danh sách xóa thông minh (listwise deletion). Cách xóa này có thể thu nhỏ đáng kể kích thước mẫu và làm cho nó cực kỳ khó khăn để phát hiện các hiệu ứng nhỏ. Do đó, một số chương trình phần mềm đưa ra những lựa chọn thay thế cho các giá trị bị thiếu bằng các giá trị ước tính thông qua một quá trình được gọi là *quy đổi* (imputation). Ví dụ, nếu giá trị bị thiếu là một mục (một chỉ số) trong một quy mô đa mục, giá trị quy đổi cho giá trị bị thiếu này có thể là mức trung bình trong số các câu trả lời hiện có của người được hỏi. Nếu giá trị thiếu thuộc về một quy mô đơn mục, nhiều nhà nghiên cứu sử dụng trung bình của các câu trả lời khác đến mục đó như là giá trị quy đổi. Khoản tính như vậy có thể bị sai lệch nếu giá trị còn thiếu có tính chất hệ thống hơn là ngẫu nhiên. Hai phương pháp có thể đưa ra những ước tính quy đổi tương đối không thiên vị là kỹ thuật tối ưu (maximum likelihood procedures) và phương pháp đa quy đổi (multiple imputation methods). Cả hai kỹ thuật này đều được hỗ trợ trong các chương trình phần mềm phổ biến như SPSS và SAS.

Chuyển đổi dữ liệu (data transformation). Đôi khi, biến đổi giá trị dữ liệu (làm sạch dữ liệu) trước khi giải thích ý nghĩa của chúng là một công việc cần thiết. Ví dụ, các chỉ số được mã hóa ngược (chúng mang nội dung trái ngược với phạm trù tương ứng) nên được đảo ngược trước khi chúng được so sánh kết hợp với các chỉ số không bị mã hóa ngược. Các loại chuyển đổi dữ liệu khác có thể bao gồm việc tạo ra các thang đánh giá bằng cách thêm các chỉ số cá nhân, tạo ra một chỉ số tổng từ một tập hợp các dữ liệu đã thu thập và

gói gọn nhiều giá trị thành các phân nhóm nhỏ hơn (ví dụ, nhóm các mức thu nhập thành một khoảng thu nhập).

Phân tích đơn biến

Phân tích đơn biến (Univariate analysis) hoặc phân tích duy nhất một biến, gồm tập hợp các kỹ thuật thống kê có thể mô tả các thuộc tính chung của một biến. Số liệu thống kê đơn biến bao gồm: (1) phân phối tần số (frequency distribution), (2) xu hướng trung tâm (central tendency), và (3) phân tán (dispersion). **Phân phối tần số** của một biến là tổng hợp mức độ thường xuyên (hoặc tỷ lệ phần trăm) của các giá trị cá biệt hoặc vùng giá trị của biến đó. Ví dụ, chúng ta có thể đo lường có bao nhiêu lần người được phỏng vấn trong mẫu tham dự các nghi lễ tôn giáo (như là một thước đo mức độ tín ngưỡng của họ) thông qua thang phân loại: không bao giờ, mỗi năm một lần, vài lần mỗi năm, khoảng mỗi tháng một lần, vài lần một tháng, vài lần mỗi tuần và một mục tùy chọn “không trả lời”. Nếu chúng ta đếm số lượng (hoặc tỷ lệ phần trăm) quan sát (câu trả lời) trong mỗi loại (ngoại trừ “không trả lời” - được coi là một giá trị thiếu chứ không phải là một mục cụ thể) và minh họa chúng trong một bảng (Hình 14.1), thì những gì chúng ta có chính là một phân phối tần số. Phân phối này cũng có thể được mô tả dưới hình thức là một biểu đồ cột, như trên bảng bên phải của hình 14.1 với trục ngang đại diện cho từng loại của biến và trục đứng đại diện cho tần số hoặc tỷ lệ phần trăm quan sát trong mỗi loại.



Hình 14.1. Phân phối tần số của biến “tín ngưỡng tôn giáo”

Đối với mẫu rất lớn có các quan sát độc lập và ngẫu nhiên, sự phân phối tần số có xu hướng đi theo đồ thị giống như một đường cong hình cái chuông (một thanh biểu đồ uyển chuyển về sự phân phối tần số) tương tự như thể hiện trong Hình 14.2 - ở đó hầu hết các quan sát được nhóm về phía trung tâm của vùng giá trị và các quan sát càng ít đi ở các đầu

cực của vùng giá trị đó. Một đường cong như vậy được gọi là một *phân phối chuẩn* (normal distribution).

Xu hướng trung tâm (Central tendency) là một ước tính về trung tâm khi phân phối các giá trị. Có ba ước tính chủ yếu của xu hướng trung tâm: *trung bình số học* (arithmetic mean), *trung vị* (median) và *yếu vị* (mode). **Trung bình số học** (thường gọi đơn giản là “trung bình”) là trung bình cộng của tất cả các giá trị trong một phân phối. Hãy xem xét một bộ tám giá trị: 15, 22, 21, 18, 36, 15, 25, 15. Trung bình cộng của các giá trị là $(15 + 20 + 21 + 20 + 36 + 15 + 25 + 15) / 8 = 20,875$. Các loại khác của giá trị trung bình bao gồm *trung bình nhân* (geometric mean - số thứ tự n của sản phẩm thứ n trong một phân phối) và *số trung bình điều hòa* (harmonic mean - nghịch đảo của trung bình số học có nghĩa là đối ứng của mỗi giá trị trong một phân phối), nhưng các giá trị trung bình này không thực sự phổ biến cho các thống kê phân tích đối với dữ liệu nghiên cứu xã hội.

Ước tính thứ hai của xu hướng trung tâm - **trung vị** - là giá trị ở giữa vùng giá trị của phân phối. Nó được tính bằng cách phân loại tất cả các giá trị trong một phân phối theo thứ tự tăng dần và chọn giá trị ở giữa. Trong trường hợp có hai giá trị ở giữa (nếu có một số chẵn của các giá trị trong phân phối), mức trung bình của hai giá trị giữa chính là trung vị. Trong ví dụ trên, các giá trị được sắp xếp là: 15, 15, 15, 18, 22, 21, 25, 36. Hai giá trị giữa 18 và 22, do đó trung vị là $(18 + 22) / 2 = 20$.

Cuối cùng, **yếu vị** là giá trị thường xuyên xuất hiện nhất trong phân phối các giá trị. Trong ví dụ trước, giá trị thường xuyên xảy ra nhất là 15, nó là yếu vị của bộ điểm số trên. Lưu ý rằng bất kỳ giá trị nào được ước tính từ một mẫu, như là giá trị trung bình, trung vị, yếu vị hoặc bất kỳ ước tính sau này đều được gọi là một **số liệu thống kê** (statistic - gọi tắt là một thống kê).

Sự phân tán (dispersion) đề cập đến mức độ các giá trị định vị quanh xu hướng trung tâm, ví dụ, các giá trị đứng rất sát nhau hay xa nhau xung quanh giá trị trung bình. Hai cách thức đo phổ biến của sự phân tán là phạm vi (khoảng biến thiên) và độ lệch chuẩn. **Phạm vi** (range) là sự khác biệt (hiệu số) giữa các giá trị cao nhất và thấp nhất trong một phân phối. Phạm vi trong ví dụ trước là $36 - 15 = 21$. Phạm vi thường đặc biệt nhạy cảm với sự hiện diện của giá trị bên ngoài. Ví dụ, nếu giá trị cao nhất trong phân phối trên 85 và các giá trị khác vẫn giữ nguyên, phạm vi là $85 - 15 = 70$. **Độ lệch chuẩn** (standard deviation) là thước đo thứ hai của sự phân tán giúp khắc phục các giá trị bên ngoài bằng cách sử dụng một công thức tính toán mức độ gần hoặc xa của mỗi giá trị từ giá trị trung

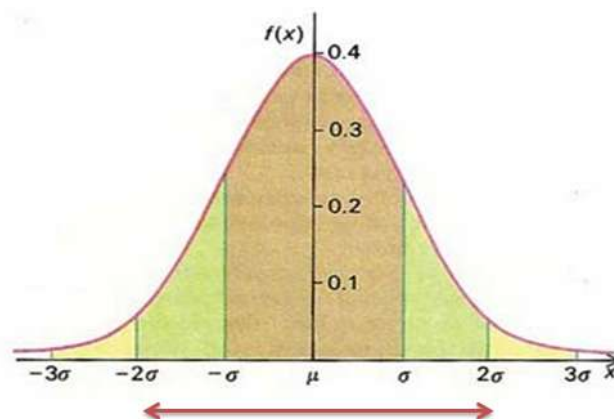
biên của phân phối:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n - 1}}$$

Trong đó σ là độ lệch chuẩn, x_i là quan sát (hoặc giá trị) thứ i , μ là trung bình số học, n là số lượng quan sát và Σ tổng số tất cả các quan sát. Bình phương của độ lệch chuẩn được gọi là **phương sai** (variance) của một phân phối. Một phân phối tần số nếu được phân phối bình thường, sẽ chứng kiến 68% của các quan sát nằm trong độ lệch chuẩn trung bình ($\mu \pm 1\sigma$), 95% của các quan sát nằm trong vòng hai lần độ lệch chuẩn ($\mu \pm 2\sigma$) và 99,7 % của các quan sát nằm trong vòng ba lần độ lệch chuẩn ($\mu \pm 3\sigma$), như tại Hình 14.2.

Phân tích lưỡng biến

Phân tích lưỡng biến (bivariate analysis) là việc xem xét xem hai biến có liên quan đến nhau như thế nào. Thống kê lưỡng biến phổ biến nhất là phân tích tương quan lưỡng biến **bivariate correlation** (thường gọi đơn giản là "tương quan"), đó là một số ở giữa -1 và +1 biểu thị sức mạnh của mối quan hệ giữa hai biến. Chẳng hạn chúng ta muốn nghiên cứu xem sự liên quan giữa tuổi tác và sự tự tin của một mẫu có 20 người, có nghĩa là, chúng ta muốn biết khi tuổi càng cao, thì sự tự tin sẽ tăng, giảm hay là không thay đổi. Nếu sự tự tin tăng lên, thì chúng ta có một mối tương quan tích cực (tương quan dương) giữa hai biến, nếu sự tự tin giảm đi, chúng ta có một mối tương quan tiêu cực (tương quan âm) và nếu nó vẫn giữ nguyên, chúng ta không có mối tương quan nào. Để tính toán giá trị của mối tương quan này, hãy xem xét các số liệu giả thuyết được thể hiện trong Bảng 14.1.



Trong phân phối chuẩn, 95% quan sát rơi vào khoảng giữa hai độ lệch chuẩn của giá trị trung bình ($\mu \pm 2\sigma$)

Hình 14.2. Phân phối chuẩn

Quan sát	Tuổi (x)	Sự tự tin (y)	xy	x ²	y ²
1	39	4.1	159.9	1521	16.81
2	45	4.6	207	2025	21.16
3	29	3.8	110.2	841	14.44
4	42	4.4	184.8	1764	19.36
5	19	3.2	60.8	361	10.24
6	22	3.1	68.2	484	9.61
7	39	3.8	148.2	1521	14.44
8	30	4.1	123	900	16.81
9	33	4.3	141.9	1089	18.49
10	23	3.7	85.1	529	13.69
11	20	3.5	70	400	12.25
12	18	3.2	57.6	324	10.24
13	24	3.7	88.8	576	13.69
14	22	3.3	72.6	484	10.89
15	29	3.4	98.6	841	11.56
16	35	4.0	140	1225	16.00
17	36	4.1	147.6	1296	16.81
18	37	3.8	140.6	1369	14.44
19	35	3.4	119	1225	11.56
20	32	3.6	115.2	1024	12.96
Tổng (Σ)	609	75.1	2339.1	19799.0	285.45

$$r_{xy} = 0.79$$

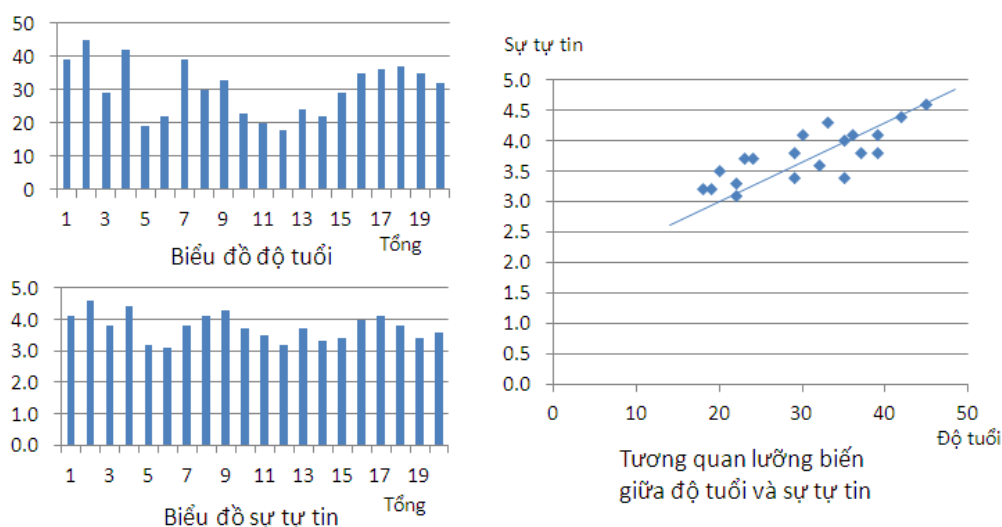
Bảng 14.1. Dữ liệu giả thuyết giữa độ tuổi và sự tự tin

Hai biến trong dữ liệu này là tuổi (x) và sự tự tin (y). Tuổi là một biến trong thang tỷ lệ, trong khi đó sự tự tin là điểm số trung bình được đo bởi thang đa mục Likert 7 điểm, chạy từ điểm “hoàn toàn không đồng ý” đến điểm “hoàn toàn đồng ý). Biểu đồ của mỗi biến được hiển thị ở phía bên trái của hình 14.3. Công thức tính toán mối tương quan 2 biến là:

$$r_{xy} = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{(n-1) s_x s_y} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}$$

Trong đó r_{xy} là mối tương quan, x và y là giá trị trung bình của biến tuổi và biến sự tự tin, s_x và s_y là độ lệch chuẩn của x và y. Giá trị tương quan giữa hai biến tuổi và sự tự tin được tính bằng công thức được minh họa trong Bảng 14.1 và bằng 0,79. Giá trị này cho thấy độ tuổi có một mối tương quan tích cực, mạnh với sự tự tin, nghĩa là sự tự tin có xu hướng tăng khi tuổi tác ngày càng tăng và giảm đi khi tuổi càng ít. Điều này cũng có thể được nhìn thấy khi trực quan so sánh độ tuổi và sự tự tin trong biểu đồ Hình 14,3, ở đó ta có thể thấy rằng những cao điểm của hai biểu đồ xuất hiện khá đều đặn cùng nhau. Lưu ý rằng các trục thẳng đứng trong Hình 14.3 đại diện cho các giá trị quan sát thực tế và không phải là tần số quan sát (như trong Hình 14.1). Do đó, đây không phải là phân phối tần số mà là biểu đồ tần số.

Biểu đồ phân tán lưỡng biến trong bảng bên phải của Hình 14.4 về bản chất là một biểu đồ của sự tự tin trên trục thẳng đứng so với độ tuổi trên trục ngang. Biểu đồ này gần giống như một đường lên dốc (độ dốc tích cực), đó cũng là dấu hiệu của một mối tương quan tích cực. Nếu hai biến có tương quan âm (tiêu cực), biểu đồ phân tán sẽ dốc xuống (độ dốc tiêu cực), ngụ ý rằng sự gia tăng tuổi tác sẽ làm suy giảm sự tự tin và ngược lại. Nếu hai biến không tương quan, biểu đồ phân tán sẽ xấp xỉ một đường ngang (không có độ dốc), nghĩa là sự gia tăng tuổi tác sẽ không có một sự liên quan có hệ thống nào với sự tự tin.



Hình 14.3. Biểu đồ tần số và biểu đồ tương quan giữa tuổi và sự tự tin

Sau khi tính toán tương quan 2 biến, các nhà nghiên cứu thường quan tâm đến việc liệu mối tương quan đó có ý nghĩa quan trọng không (ý nghĩa thực tế) hoặc nó đơn thuần gây ra bởi một sự tình cờ ngẫu nhiên. Trả lời một câu hỏi như vậy sẽ yêu cầu kiểm tra giả thuyết sau đây:

$$H_0: r = 0$$

$$H_1: r \neq 0$$

H_0 được gọi là *giả thuyết vô hiệu* (null hypothesis), H_1 được gọi là *giả thuyết thay thế hay đối thuyết* (alternative hypothesis), đôi khi được kí hiệu là H_a). Mặc dù chúng có vẻ giống như hai giả thuyết, nhưng H_0 và H_1 thực sự đại diện cho một giả thuyết duy nhất và đối lập trực tiếp với nhau hoàn toàn. Chúng ta quan tâm đến thử nghiệm H_1 hơn là H_0 . Cũng lưu ý rằng H_1 là một giả thuyết không định hướng vì nó không chỉ định cho dù r là lớn hơn hoặc nhỏ hơn không. Giả thuyết định hướng sẽ được quy định cụ thể như $H_0: r \leq 0$; $H_1: r > 0$ (nếu chúng ta đang thử nghiệm cho một mối tương quan tích cực). Việc thử

thử nghiệm ý nghĩa của các giả thuyết định hướng được thực hiện bằng cách sử dụng *kiểm định t một đầu* (one-tailed t-test), trong khi đó thử nghiệm ý nghĩa của các giả thuyết không định hướng được thực hiện bằng cách sử dụng *kiểm định t hai đầu* (two-tailed t-test).

Trong kiểm định thống kê, đối thuyết không kiểm tra trực tiếp được. Thay vào đó, nó được thử nghiệm gián tiếp bằng cách loại bỏ các giả thuyết vô hiệu với một mức độ xác suất nhất định. Thử nghiệm thống kê luôn luôn có tính xác suất, bởi vì chúng ta không bao giờ có thể khẳng định chắc chắn 100% khi mà suy luận của chúng ta dựa trên số liệu mẫu nhưng lại áp dụng cho toàn bộ quần thể nghiên cứu, trong khi đó mẫu bao giờ cũng nhỏ hơn quần thể. Xác suất của một suy luận thống kê được gọi là *giá trị p* (p-value). Giá trị p được so sánh với *mức ý nghĩa* (significance level) (α). Mức ý nghĩa là mức rủi ro tối đa khi các suy luận có thể không chính xác nhưng chúng ta có thể chấp nhận được. Đối với hầu hết các phân tích thống kê, α được thiết lập đến 0,05. Nếu giá trị p nhỏ hơn $\alpha = 0,05$ thì ta có đủ bằng chứng thống kê để bác bỏ giả thuyết vô hiệu và do đó, gián tiếp chấp nhận giả thuyết thay thế. Nếu $p > 0,05$, chúng ta không có đủ bằng chứng thống kê để bác bỏ giả thuyết vô hiệu hoặc chấp nhận giả thuyết thay thế.

Cách đơn giản nhất để kiểm tra giả thuyết trên là tìm kiếm các giá trị tới hạn của r được liệt kê trong các bảng thống kê có sẵn trong tất cả sách khoa học thống kê hay có thể tra cứu trên mạng Internet (hầu hết các chương trình phần mềm cũng cung cấp các thử nghiệm về mức ý nghĩa). Giá trị tới hạn r phụ thuộc vào mức ý nghĩa mong muốn ($\alpha = 0,05$), bậc tự do (df) và phụ thuộc vào liệu kiểm định dự kiến là kiểm định một đầu hay hai đầu. **Bậc tự do** (degree of freedom) là số lượng các giá trị có thể thay đổi tự do trong bất kỳ tính toán thống kê nào. Trong trường hợp có sự tương quan, df chỉ đơn giản bằng $n - 2$ hoặc đối với các dữ liệu trong Bảng 14.1, df là $20 - 2 = 18$.

Có hai bảng thống kê khác nhau cho thử nghiệm một phía và hai phía. Trong bảng hai phía, hệ số tương quan r cho $\alpha = 0,05$ và $df = 18$ là 0,44. Để cho tương quan của chúng ta (bằng 0,79) trở nên có ý nghĩa, nó phải lớn hơn giá trị tới hạn 0,44 hoặc ít hơn -0,44. Bởi vì 0,79 lớn hơn 0,44, chúng ta kết luận rằng có một mối tương quan đáng kể (có ý nghĩa) giữa tuổi và sự tự tin trong tập dữ liệu của chúng ta, hay nói cách khác, sự chênh lệch là ít hơn 5% nên sự tương quan này là một khả năng có thể xảy ra. Vì vậy, chúng ta có thể từ chối giả thuyết vô hiệu với $r \leq 0$ và phát biểu một cách gián tiếp rằng giả thuyết thay thế $r > 0$ gần như là đúng.

Hầu hết các nghiên cứu có nhiều hơn hai biến. Nếu có n biến thì chúng ta sẽ có tổng cộng $n*(n-1)/2$ mối tương quan có thể có giữa n biến. Mỗi tương quan như vậy được dễ dàng tính toán bằng cách sử dụng một chương trình phần mềm như SPSS, chứ không phải tính toán thủ công bằng cách sử dụng công thức cho sự tương quan (như đã làm trong Bảng 14.1), chúng được thể hiện bằng cách sử dụng một mô hình tương quan, như trong Bảng 14.2. Mô hình tương quan liệt kê các tên biến ở các hàng và cột đầu tiên, sau đó mô tả mối tương quan lưỡng biến giữa các cặp của các biến tại các ô thích hợp trong mô hình. Các giá trị nằm trên đường chéo chính (từ góc trên cùng bên trái sang góc dưới cùng bên phải) của mô hình này luôn luôn là 1, bởi vì bất kỳ biến nào cũng luôn luôn có tương quan tuyệt đối với chính nó. Hơn nữa, đối với mỗi tương quan không định hướng, mỗi tương quan giữa các biến số $V1$ và $V2$ là giống như giữa $V2$ và $V1$. Do đó, mô hình tam giác thấp hơn (giá trị dưới đường chéo chính) là một sự phản chiếu của mô hình tam giác trên (giá trị trên đường chéo chính), vì vậy để đơn giản, chúng ta thường chỉ liệt kê mô hình tam giác dưới. Nếu các mối tương quan chứa các biến được đo bằng thang khoảng thì loại hình cụ thể của mối tương quan này được gọi là *mối tương quan hồi quy Pearson* (Pearson product moment correlations).

Một cách hữu ích khác để trình bày dữ liệu lưỡng biến là **kỹ thuật phân tích bảng chéo** (cross-tabulation) (thường được viết tắt là cross-tab và đôi khi được gọi chính thức là một *bảng ngẫu nhiên* - contingency table). Bảng chéo là một bảng mô tả tần số (hoặc tỷ lệ phần trăm) của tất cả bộ kết hợp của hai hay nhiều biến định danh hoặc phân loại. Ví dụ, hãy giả sử rằng chúng ta có những quan sát sau đây cho giới tính và điểm số của một mẫu 20 sinh viên, như thể hiện trong Bảng 14.2. Giới tính là một biến định danh (nam/nữ hoặc M/F) và điểm số là một biến phân loại với ba cấp độ (A, B và C). Một bảng dữ liệu chéo đơn giản hiển thị sự phân phối chung của giới tính và điểm số (tức là, bao nhiêu học sinh của mỗi loại giới tính nằm trong mỗi loại điểm số, nó như một tần số thực hoặc như là một tỷ lệ phần trăm) trong một ma trận 2×3 . Ma trận này sẽ giúp chúng ta xem xét A, B và C là điểm phân phối cân bằng của sinh viên nam và nữ. Các dữ liệu chéo trong Bảng 14.3 cho thấy sự phân phối điểm A có xu hướng nghiêng về các nữ sinh: trong một mẫu có 10 nam và 10 nữ, 5 sinh viên nữ so với duy nhất một sinh viên nam nhận điểm A. Ngược lại, sự phân phối điểm C lại nghiêng về phía các nam sinh viên: 3 nam sinh đã nhận được điểm C, so với chỉ 1 nữ sinh viên có điểm C. Ngoài ra, sự phân phối của điểm B là tương đối đồng đều, với 6 sinh viên nam và 5 sinh viên nữ. Dòng cuối cùng và cột cuối cùng của bảng này được gọi là tổng biên độ bởi vì nó cho biết tổng số trên mỗi thể loại và hiển thị dọc theo rìa

của bảng.

	V1	V2	V3	V4	V5	V6	V7	V8
V1	1.000							
V2	0.274	1.000						
V3	-0.134	-0.269	1.000					
V4	0.201	-0.153	0.075	1.000				
V5	-0.095	-0.166	0.278	-0.011	1.000			
V6	-0.129	0.280	-0.348	-0.378	-0.009	1.000		
V7	0.171	-0.122	0.296	0.086	0.193	0.233	1.000	
V8	0.518	0.238	0.238	-0.227	-0.551	0.082	-0.102	1.000

Bảng 14.2. Ma trận giả thuyết tương quan cho tám biến

Quan sát	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Giới tính	F	M	F	M	F	M	M	M	F	F	M	M	M	F	F	M	F	F	F	M
Mức độ	A	B	B	B	C	A	C	B	B	A	B	C	C	B	B	B	A	A	B	B

Một tập hợp dữ liệu giả thuyết

		Mức độ			Tổng	
			A	B	C	
Giới tính	Nam	Tổng số	1	6	3	10
		Tổng số ước đạt	2.5	5.5	2.0	
	Nữ	Tổng số	4	5	1	10
		Tổng số ước đạt	2.5	5.5	2.0	
Tổng			5	11	4	20

Bảng phân tích chéo giữa giới tính và độ tuổi

Tổng
cận biên

Bảng 14.3. Ví dụ về phân tích bảng chéo

Mặc dù chúng ta có thể thấy một mô hình khác biệt về phân phối điểm số giữa sinh viên nam và nữ trong Bảng 14.3, nhưng nó là một mô hình thực tế hay "có ý nghĩa về thống kê" hay không? Nói cách khác, các giá trị tần số trên có phản ánh đúng những gì chúng ta dự kiến trên thực tế? Để trả lời câu hỏi này, chúng ta nên tính toán giá trị dự kiến của các quan sát trong mỗi ô của ma trận chéo 2x3. Nó được thực hiện bằng cách nhân tổng số cột biên và số hàng tổng biên của mỗi ô và chia cho tổng số quan sát. Ví dụ, ở cặp ô nam/điểm, tính toán mong đợi = $5 \cdot 10 / 20 = 2,5$. Nói cách khác, chúng ta đã dự kiến 2,5 nam sinh viên nhận được điểm A, nhưng trong thực tế, chỉ có một nam sinh nhận được điểm A. Để xác định sự khác biệt giữa giá trị dự kiến và giá trị thực tế là đáng kể hay không đáng kể, chúng ta có thể sử dụng *kiểm định Chi bình phương* (chi-square test). Các số liệu thống kê chi-square được tính là sự khác biệt trung bình giữa các giá trị đã quan sát

và giá trị dự kiến dọc theo tất cả các ô. Chúng ta có thể so sánh giá trị này với giá trị tới hạn trong một mức xác suất mong muốn ($p < 0,05$) và bậc tự do, nó chỉ đơn giản là $(m-1)*(n-1)$, trong đó m và n là số lượng hàng và cột tương ứng. Trong ví dụ này, $df = (2-1)*(3-1) = 2$. Từ bảng tiêu chuẩn chi-square trong bất kỳ cuốn sách thống kê nào, giá trị tới hạn chi-square cho $p = 0,05$ và $df = 2$ là 5,99. Các tính toán giá trị chi-square dựa trên các dữ liệu quan sát là 1.00, thấp hơn giá trị tới hạn. Do đó, chúng ta phải kết luận rằng dưới góc độ thống kê, mô hình về điểm số đã quan sát không khác so với mô hình trên thực tế.

Chương 15

PHÂN TÍCH ĐỊNH LƯỢNG: THỐNG KÊ SUY DIỄN

Thống kê suy diễn (thống kê suy luận - inferential statistics) là các phương thức thống kê được sử dụng để đưa ra các kết luận về mối liên hệ giữa các biến. Khác với các thống kê mô tả, thống kê suy diễn được thiết kế một cách rõ ràng để kiểm định giả thuyết. Có rất nhiều phương pháp thống kê được áp dụng trong hình thức này, hầu hết trong số đó được hỗ trợ bởi các phần mềm thống kê hiện đại như SPSS và SAS. Chương này sẽ cung cấp một bức tranh tổng quát nhất về những khái niệm và các kỹ thuật quan trọng. Để tìm hiểu sâu thêm về các nội dung liên quan, độc giả nên tham khảo thêm các tài liệu thống kê hoặc tham gia một khóa học về số liệu thống kê.

Khái niệm cơ bản

Nhà triết học người Anh Karl Popper nói rằng chúng ta không bao giờ có khả năng chứng minh các lý thuyết, mà chỉ có thể bác bỏ chúng. Ví dụ, làm thế nào chúng ta có thể chứng minh rằng mặt trời sẽ mọc vào ngày mai? Popper nói rằng chỉ vì trước tới giờ, mặt trời luôn mọc hàng ngày nên chúng ta có thể suy ra nó sẽ mọc vào ngày mai. Nhưng điều đó không có nghĩa chắc chắn ngày mai mặt trời sẽ mọc, bởi vì lý luận quy nạp như vậy chỉ là phỏng đoán và một hiện tượng trong tương lai có thể có hoặc có thể không xảy ra. Thay vào đó, ông cho rằng chúng ta có thể giả định một lý thuyết rằng mặt trời sẽ mọc mỗi ngày mà không nhất thiết phải chứng minh lý thuyết đó. Và nếu mặt trời không mọc vào một ngày bất kỳ nào đó, lý thuyết này không đúng và bị bác bỏ.

Tương tự như vậy, chúng ta chỉ có thể bác bỏ giả thuyết dựa trên những bằng chứng đối lập, nhưng chúng ta không bao giờ có thể thực sự chấp nhận chúng bởi vì sự hiện diện của bằng chứng hiện tại không có nghĩa là trong tương lai chúng ta không tìm thấy những bằng chứng đối lập. Bởi vì chúng ta không thể thực sự chấp nhận một giả thuyết thay thế (*đối thuyết*) nên chúng ta phải hình thành một *giả thuyết vô hiệu* (null hypothesis) ngược lại với đối thuyết thay thế. Sau đó chúng ta sử dụng bằng chứng thực nghiệm để bác bỏ giả thuyết không; bằng cách này gián tiếp chứng minh và bảo vệ cho các đối thuyết thay thế ở một mức độ xác suất nào đó.

Vấn đề thứ hai liên quan tới việc kiểm tra mối quan hệ của các giả thuyết trong

nghiên cứu khoa học xã hội là biến phụ thuộc có thể bị ảnh hưởng bởi một số lượng rất nhiều các biến không liên quan và việc kiểm soát tất cả các ảnh hưởng không mong muốn là điều khó có thể đạt được. Do đó, ngay cả khi hai biến có vẻ có liên quan trong một mẫu quan sát, nó có thể không thực sự liên quan tới toàn bộ khách thể nghiên cứu và do đó số liệu suy luận thống kê là không bao giờ chắc chắn hoặc cố định, mà luôn luôn có một mức độ xác suất nào đó.

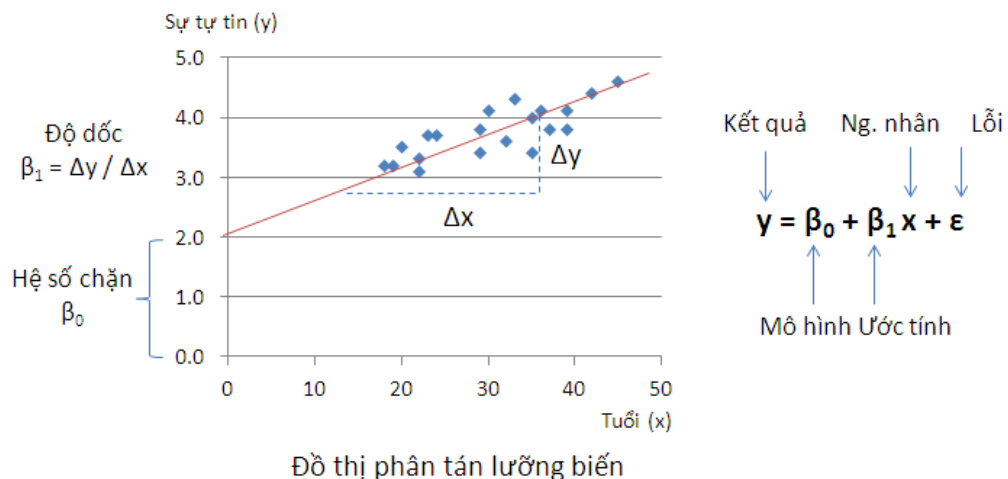
Làm thế nào để biết liệu một mối quan hệ giữa hai biến số trong một mẫu quan sát là có ý nghĩa, có giá trị chứ không phải hoàn toàn ngẫu nhiên? Ronald A. Fisher - một trong những nhà thống kê nổi tiếng nhất trong lịch sử, đã thiết lập ra những định hướng cơ bản để thử nghiệm mức độ có ý nghĩa này. Ông cho rằng một kết quả thống kê có thể được coi là đáng kể (có ý nghĩa, có giá trị, chấp nhận được) nếu nó chỉ ra được rằng xác suất bị bác bỏ của nó (do sự ngẫu nhiên mang đến) là nhỏ hơn hoặc bằng 5%. Trong suy luận thống kê, xác suất này được gọi là *giá trị xác suất p* (p-value), 5% được gọi là *mức ý nghĩa* (α), mức độ mong muốn trong quan hệ giữa p-value và α được ký hiệu là: $p \leq 0,05$. *Mức ý nghĩa* là mức độ rủi ro tối đa mà chúng ta sẵn sàng chấp nhận, được coi là mức giá trị của những suy luận rút ra mẫu áp dụng cho toàn bộ khách thể nghiên cứu. Nếu giá trị p nhỏ hơn 0,05 hoặc 5% thì nó có nghĩa rằng chúng ta có 5% cơ hội là không chính xác trong việc bác bỏ giả thuyết vô hiệu và được liệt vào Lỗi loại I (Type I error). Nếu $p > 0,05$, chúng ta không có đủ bằng chứng để bác bỏ giả thuyết vô hiệu hay chấp nhận giả thuyết thay thế.

Chúng ta cũng phải hiểu rõ ba khái niệm liên quan đến thống kê: *phân phối mẫu*, *sai số chuẩn* và *khoảng tin cậy*. **Phân phối mẫu** (sampling distribution) là sự phân phối lý thuyết của một số lượng vô hạn các mẫu lấy từ khách thể nghiên cứu. Tuy nhiên, bởi vì bất kỳ mẫu nào cũng không bao giờ có thể giống hệt với khách thể nghiên cứu nên trong mỗi một mẫu luôn luôn có một mức độ khác biệt so với khách thể nghiên cứu, nó được gọi là **sai số chuẩn** (standard error). Nếu sai số tiêu chuẩn này nhỏ thì những ước tính thống kê bắt nguồn từ mẫu (ví dụ giá trị trung bình của mẫu) sẽ là ước tính tương đối chính xác cho toàn bộ khách thể nghiên cứu. Mức độ chính xác của những ước tính trong mẫu được định nghĩa là một **khoảng tin cậy** (confidence interval - CI). Khoảng CI 95% được xác định là một khoảng giá trị cộng hoặc trừ hai lần độ lệch chuẩn của các ước tính trung bình khi chúng được rút ra từ các mẫu khác nhau trong một phân phối mẫu. Do đó, khi chúng ta nói rằng số liệu thống kê trong mẫu nghiên cứu của chúng ta có **kháng CI là 95%** có nghĩa là chúng ta có thể tự tin nói rằng với 95% xác suất chính xác, các thông số của khách thể

nghiên cứu nằm trong khoảng hai lần độ lệch chuẩn của các thông tin trong mẫu. Cùng với nhau, giá trị xác suất p và khoảng tin cậy CI cung cấp cho chúng ta một ý tưởng tốt về xác suất chính xác trong kết quả nghiên cứu của chúng ta thông qua mức độ tương ứng của chúng với các thông số của khách thể nghiên cứu.

Mô hình tuyến tính tổng quát

Hầu hết các phương pháp suy luận thống kê trong nghiên cứu khoa học xã hội có nguồn gốc chung từ mô hình tuyến tính tổng quát (General Linear Model - GLM). *Mô hình* là một phương trình toán học ước tính có thể được sử dụng để đại diện cho một tập hợp các dữ liệu và *tuyến tính* liên quan đến một đường thẳng. Do đó, GLM là một hệ thống các phương trình biểu thị bằng những mô hình tuyến tính có thể được sử dụng để đại diện cho các mối quan hệ xuất hiện từ dữ liệu đã thu thập.



Hình 15.1. Mô hình tuyến tính hai biến

Loại đơn giản nhất của GLM là một mô hình tuyến tính hai biến nhằm xem xét mối quan hệ giữa một biến độc lập (nguyên nhân hay đầu vào hay yếu tố đầu vào) và một biến phụ thuộc (kết quả hoặc đầu ra). Giả định rằng chúng ta có hai biến tương ứng là tuổi tác và sự tự tin. Đồ thị phân tán lưỡng biến biểu thị mối quan hệ giữa hai biến này được thể hiện trong Hình 15.1, trong đó tuổi (yếu tố dự đoán) nằm ngang trục x và sự tự tin (kết quả) nằm dọc trục y. Đồ thị scatterplot cho thấy các quan sát độc lập đại diện cho sự kết hợp của tuổi tác và sự tự tin dường như nằm rải rác xung quanh một đường thẳng lên dốc. Chúng ta có thể ước tính các thông số của đường này, chẳng hạn như độ dốc và điểm chặn của GLM. Như chúng ta đã học môn đại số thời trung học, nhớ lại rằng đường thẳng có thể

được đại diện bằng cách sử dụng các phương trình toán học $y = mx + c$, ở đó m là *độ dốc* (*slope*) của đường thẳng (số lượng thay đổi của y khi x thay đổi 1 đơn vị) và c là *hệ số hạng chặn* (*intercept*) (là giá trị của y khi x bằng 0). Trong GLM, phương trình này được đại diện chính thức là:

$$y = \beta_0 + \beta_1 x + \varepsilon$$

trong đó β_0 là độ dốc, β_1 là số hạng chặn và ε là *sai số*. ε đại diện cho độ lệch của các quan sát thực tế so với các giá trị ước tính của nó, vì hầu hết các quan sát chỉ nằm gần với đường thẳng chứ không nằm chính xác trên đường thẳng đó (xem đường GLM). Lưu ý rằng một mô hình tuyến tính có thể có nhiều hơn hai yếu tố dự báo. Để hình dung một mô hình tuyến tính với hai dự đoán, hãy tưởng tượng một hình khối có ba chiều, với kết quả (y) theo trục thẳng đứng và hai yếu tố dự báo (x_1 và x_2) dọc theo hai trục ngang của khối ba chiều. Một đường thẳng mô tả mối quan hệ giữa hai hay nhiều biến được gọi là một đường hồi quy, β_0 và β_1 (và các giá trị beta khác) được gọi là hệ số hồi quy và quá trình ước lượng hệ số hồi quy được gọi là *phân tích hồi quy*. GLM dùng cho phân tích hồi quy với n biến dự đoán là:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n + \varepsilon$$

Trong phương trình trên, các biến dự báo x_i có thể đại diện cho các biến độc lập hoặc đồng biến (biến đối chứng). Đồng biến không phải là các biến trọng tâm trong phân tích nhưng chúng có thể có một số tác động vào biến phụ thuộc y và vì vậy phải được kiểm soát, do đó những ảnh hưởng đối với các biến độc lập sẽ được phát hiện chính xác hơn. Đồng biến giúp phát hiện các lỗi hệ thống trong một phương trình hồi quy, trong khi đó sai số (ε) cho phép phát hiện các lỗi ngẫu nhiên.

Mặc dù hầu hết các biến trong mô hình GLM thường là các biến khoảng hoặc biến tỷ lệ nhưng cũng có những ngoại lệ. Một số các biến dự đoán thậm chí có thể trở thành các biến định danh (ví dụ, giới tính: nam hay nữ), được mã hóa như là *biến giả* (*dummy variable*). Đây là những biến số có thể giả định có một trong hai giá trị: 0 hoặc 1 (trong ví dụ giới tính, "nam" có thể được chỉ định là 0 và "nữ" là 1 hoặc ngược lại). Một tập hợp n biến định danh được miêu tả bằng cách sử dụng $n-1$ biến giả. Ví dụ, lĩnh vực kinh tế bao gồm nông nghiệp, công nghiệp và dịch vụ có thể được miêu tả bằng cách kết hợp hai biến giả (x_1, x_2) trong đó (0, 0) cho nông nghiệp, (0, 1) cho công nghiệp, và (1, 1) cho dịch vụ. Các biến định danh được mã hóa là 0 hay là 1 không có ý nghĩa số học là 1 lớn hơn 0 mà việc đánh số chỉ là để phục vụ các phân tích của máy tính.

GLM là một công cụ thống kê rất mạnh bởi vì nó không phải là một phương pháp thống kê đơn giản mà là một tập hợp các phương pháp có thể được sử dụng để tiến hành các phân tích phức tạp với nhiều biến dự đoán, biến kết quả và với nhiều cách thức phân tích khác nhau. Nếu chúng ta có một biến dự đoán giả và chúng ta đang so sánh ảnh hưởng của hai mức (0 và 1) trong biến giả này tới biến kết quả, tức là chúng ta đang thực hiện một *phân tích phương sai* (analysis of variance - ANOVA). Nếu chúng ta thực hiện một phân tích ANOVA đồng thời với việc kiểm soát các tác động của một hoặc nhiều đồng biến thì chúng ta có một phân tích hiệp biến (analysis of covariance - ANCOVA). Chúng ta cũng có thể có nhiều biến kết quả (ví dụ, y_1, y_1, \dots, y_n), khi đó chúng được mô tả bằng cách sử dụng “hệ phương trình” (system of equations), trong đó mỗi phương trình biểu thị cho một biến kết quả (mỗi phương trình sẽ có một bộ hệ số hồi quy riêng).

Nếu nhiều biến kết quả được mô hình hóa bằng cách dự đoán từ cùng một tập hợp biến dự báo thì việc phân tích kết quả đó được gọi là *hồi quy đa biến* (multiple regression). Nếu chúng ta thực hiện phân tích ANOVA hoặc ANCOVA với nhiều biến kết quả thì việc phân tích kết quả được lần lượt gọi là *ANOVA đa biến* (MANOVA) hoặc *ANCOVA đa biến* (MANCOVA). Nếu ta mô hình hóa kết quả trong một phương trình hồi quy mà nó được xem là một yếu tố dự báo trong một phương trình khác trong hệ thống các phương trình hồi quy có liên quan thì chúng ta có một kiểu phân tích rất phức tạp gọi là *mô hình phương trình cấu trúc* (structural equation modeling). Vấn đề thu hút sự quan tâm nhất trong GLM là quá trình *thiết lập mô hình* (model specification), nghĩa là, làm thế nào để xác định một phương trình (hoặc một hệ phương trình) hồi quy đã đưa ra mô tả hoàn hảo nhất về hiện tượng quan tâm. Việc thiết lập mô hình phải trên những tính toán lý thuyết liên qua tới hiện tượng nghiên cứu, đồng thời phải tránh việc cố tình thiết lập mô hình sao cho phù hợp với dữ liệu quan sát. Dữ liệu có vai trò củng cố chứ không phải để thiết lập mô hình.

So sánh hai nhóm

Một trong những phân tích suy luận đơn giản nhất là so sánh các kết quả hậu thử nghiệm của các đối tượng thuộc nhóm thử nghiệm và nhóm đối chứng trong một thiết kế nhóm đối chứng ngẫu nhiên, chẳng hạn như liệu sinh viên ghi danh vào một chương trình giảng dạy toán học cải cách có kết quả học tập tốt hơn so với những học viên tham gia chương trình giảng dạy toán học truyền thống. Trong trường hợp này, biến dự báo là một biến giả (1 = nhóm thử nghiệm, 0 = nhóm đối chứng) và biến đầu ra - kết quả học tập - là biến trong thang tỷ lệ (là điểm số của một bài kiểm tra toán sau khi áp dụng chương trình cải cách). Kỹ thuật phân tích cho thiết kế đơn giản này là ANOVA một chiều (một chiều

bởi vì nó chỉ liên quan đến một biến dự đoán) và việc kiểm định thống kê được sử dụng được gọi là *Student's t-test* (kiểm định t của sinh viên - gọi tắt là t-test).

T-test được giới thiệu lần đầu vào năm 1908 bởi nhà hóa học William Sealy Gosset, người chuyên giám sát chất lượng của bia đen cho nhà máy bia Guinness tại Dublin, Ireland - bia đen là loại bia phổ biến với các nhân viên khuân vác trong thế kỷ XIX ở London. Do ông chủ của Gosset không muốn tiết lộ các kỹ thuật thống kê dùng để kiểm soát chất lượng bia, Gosset đã giấu ông chủ xuất bản kỹ thuật này trong Tạp chí *Biometrika* với bút danh "Student" (ông là học trò của Sir Ronald Fisher). Phương pháp này chứa đựng những tính toán về giá trị t, một ký hiệu được Fisher sử dụng thường xuyên để biểu thị sự khác biệt giữa hai nhóm. Do đó, cái tên *Kiểm định t* với bút danh "Student" được biết đến một cách thân mật với nhà thống kê này.

T-test xem xét liệu rằng giá trị trung bình của hai nhóm có khác nhau về mặt thống kê hay không (thử nghiệm không định hướng - thử nghiệm hai phía) hay là một nhóm có trung bình thống kê lớn hơn (hoặc nhỏ hơn) nhóm khác (thử nghiệm có định hướng - thử nghiệm một phía). Trong ví dụ ở trên, nếu chúng ta muốn xem xét liệu học sinh trong chương trình giảng dạy toán cải cách học tốt hơn so với những học sinh trong chương trình giảng dạy truyền thống, vậy thì chúng ta có một thử nghiệm một phía. Giả thuyết này có thể được thiết lập như sau:

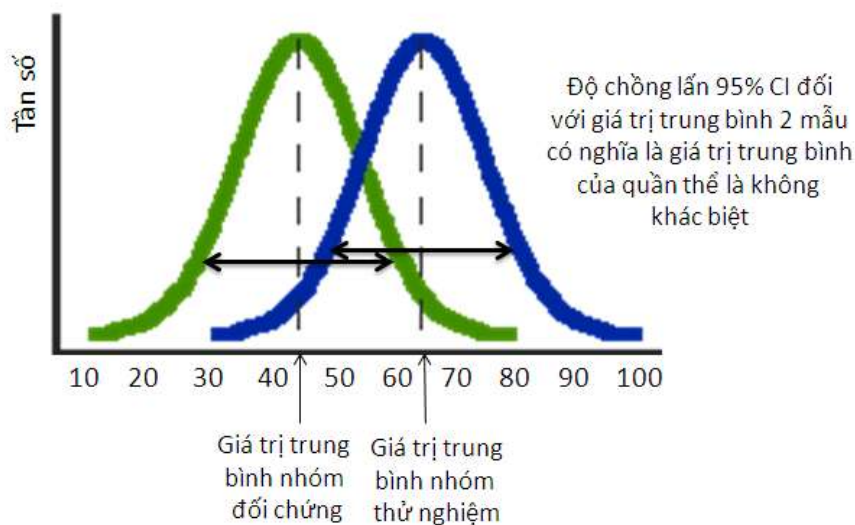
$H_0: \mu_1 \leq \mu_2$ (null hypothesis - giả thuyết vô hiệu hay giả thuyết không)

$H_1: \mu_1 > \mu_2$ (alternative hypothesis - giả thuyết thay thế hay đối thuyết)

trong đó μ_1 đại diện cho giá trị trung bình của quần thể nghiên cứu là nhóm học sinh tham gia vào chương trình cải cách (nhóm thử nghiệm) và μ_2 là giá trị trung bình của nhóm học sinh trong chương trình truyền thống (nhóm đối chứng). Lưu ý rằng các *giả thuyết vô hiệu* luôn luôn có ký hiệu "bằng" và mục tiêu của tất cả các thử nghiệm thống kê là để bác bỏ giả thuyết không.

Làm thế nào chúng ta có thể xác định sự khác biệt trong giá trị trung bình của khách thể nghiên cứu với giá trị trung bình của mẫu? Từ các giả thuyết về giá trị phân bố tần số của nhóm thực nghiệm và nhóm đối chứng trong Hình 15.2, nhóm đối chứng là một phân phối hình chuông (phân phối chuẩn) với giá trị trung bình là 45 (trên thang điểm từ 0-100), trong khi đó nhóm thử nghiệm có điểm trung bình là 65. Các giá trị trung bình này có vẻ khác nhau, nhưng chúng đều chỉ là giá trị trung bình của mẫu (\bar{X}) và chúng có thể khác với giá trị trung bình của khách thể nghiên cứu tương ứng (μ) do những sai số trong lúc chọn

mẫu. Trung bình mẫu là ước tính xác suất cho giá trị trung bình của khách thể nghiên cứu trong giới hạn một khoảng tin cậy nhất định (CI). 95% CI là giá trị trung bình của mẫu cộng trừ hai lần sai số chuẩn, trong đó sai số chuẩn là độ lệch chuẩn của phân phối trong trung bình của mẫu, được lấy từ số lượng vô hạn các mẫu của khách thể nghiên cứu. Do đó, giá trị trung bình của khách thể nghiên cứu không chỉ phụ thuộc vào giá trị trung bình mẫu mà còn phụ thuộc vào sai số chuẩn hoặc mức độ lây lan trong việc phân phối tần số của trung bình mẫu. Nếu mức độ lây lan lớn (ví dụ, hai đường cong hình chuông có rất nhiều điểm trùng nhau), thì khoảng tin cậy CI = 95% của hai giá trị trung bình cũng có thể chồng lên nhau. Như vậy, chúng ta không dám tự tin khẳng định rằng giá trị trung bình của khách thể nghiên cứu tương ứng là khác nhau đáng kể ($p < 0.05$). Ngược lại, nếu các đường cong có mức độ lây lan hẹp (ví dụ, chúng ít chồng lấn lên nhau) thì khoảng tin cậy CI của mỗi giá trị trung bình có thể không chồng lên nhau và do đó, chúng ta bác bỏ giả thuyết vô hiệu và khẳng định rằng giá trị trung bình của khách thể nghiên cứu của hai nhóm là khác nhau đáng kể tại $p < 0,05$.



Hình 15.2. Kiểm định t (t-test)

Để thực hiện t-test, đầu tiên chúng ta phải tính toán một thống kê t về sự khác biệt, đó chính là trung bình mẫu giữa hai nhóm. Thống kê này là tỷ lệ khác biệt của các trung bình mẫu có liên quan đến sự khác biệt về điểm số của chúng (sai số chuẩn):

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1 - \bar{X}_2}}$$

trong đó, tử số là sự khác biệt trong các trung bình mẫu giữa nhóm thử nghiệm

(nhóm 1) và nhóm đối chứng (nhóm 2) và mẫu số là sai số chuẩn giữa hai nhóm; do đó, có thể được ước tính như sau:

$$s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

s_2 là phương sai và n là kích thước mẫu của mỗi nhóm. Thống kê t sẽ dương nếu giá trị trung bình của nhóm thử nghiệm lớn hơn so với giá trị trung bình của nhóm đối chứng. Để kiểm tra nếu thống kê t này là đủ lớn hay chưa, chúng ta phải tìm giá trị xác suất p liên quan tới giá trị thống kê t được tính toán trong các bảng thống kê. Những bảng thống kê này có sẵn ở trong các sách về thống kê, trên mạng Internet hoặc trong chương trình phần mềm thống kê như SAS và SPSS. Bậc tự do của các số liệu thống kê t được tính như sau:

$$d.f. = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1 - 1) + (s_2^2/n_2)^2/(n_2 - 1)}$$

giá trị này thường xấp xỉ bằng (n_1+n_2-2) . Nếu giá trị xác suất p này nhỏ hơn *mức ý nghĩa* mong muốn ($\alpha = 0,05$) hoặc mức độ rủi ro cao nhất (mức xác suất) là chấp nhận được để có thể đưa ra kết luận rằng thử nghiệm là có hiệu quả (nằm trong Lỗi loại I), thì chúng ta có thể bác bỏ giả thuyết không.

Sau khi chứng minh được rằng nhóm thử nghiệm có giá trị trung bình cao hơn đáng kể so với nhóm đối chứng, câu hỏi tiếp theo thường là *mức độ ảnh hưởng* (effect size - ES) hay còn gọi là độ chênh lệch giữa nhóm thử nghiệm và nhóm đối chứng là bao nhiêu? Chúng ta có thể ước tính mức độ ảnh hưởng ES bằng cách tiến hành phân tích hồi quy trong đó mức độ ảnh hưởng là các biến kết quả (y) và biến giả thử nghiệm được mã hóa là biến dự đoán (x) trong một mô hình tuyến tính tổng quát (GLM) lưỡng biến. Hệ số hồi quy của biến đối chứng (β_1), nó cũng là độ dốc của đường hồi quy ($\beta_1 = \Delta y / \Delta x$), là một ước tính về mức độ ảnh hưởng EZ. Trong ví dụ trên, vì x là một biến giả với hai giá trị (0 và 1), $\Delta x = 1-0 = 1$, do đó mức độ ảnh hưởng hoặc β_1 chỉ đơn giản là sự khác biệt giữa các giá trị trung bình của nhóm thử nghiệm và đối chứng ($\Delta y = y_1 - y_2$).

Các thiết kế giai thừa

Mở rộng từ ví dụ trước, chúng ta hãy giả sử rằng tác động của chương trình dạy toán cải cách (thử nghiệm) đối với chương trình giảng dạy truyền thống (đối chứng) phụ thuộc vào thời lượng giảng dạy (3 hoặc 6 giờ/tuần). Bây giờ, chúng ta có thiết kế giai thừa 2x2,

với hai yếu tố là phương pháp giảng dạy (cải cách hay truyền thống) và thời lượng giảng dạy (3 hay 6 giờ/tuần). Thiết kế này không chỉ giúp chúng ta ước tính ảnh hưởng riêng lẻ của từng yếu tố, được gọi là *ảnh hưởng chính* (main effect), mà còn xác định được ảnh hưởng chung của cả hai yếu tố này, được gọi là *ảnh hưởng tương tác* (interaction effect). Mô hình tuyến tính tổng quát cho thiết kế thừa số hai chiều này được xác định như sau:

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1x_2 + \varepsilon$$

Trong đó, y đại diện cho điểm học tập của sinh viên sau thử nghiệm, x_1 là phương pháp thử nghiệm (chương trình giảng dạy cải cách hay truyền thống), x_2 thời lượng giảng dạy (3 hay 6 giờ/tuần). Lưu ý rằng cả x_1 và x_2 đều là biến giả và mặc dù x_2 trông giống như một biến trong thang tỷ lệ (3 hoặc 6) nhưng nó thực sự đại diện cho hai phân nhóm trong thiết kế thừa số. Hệ số hồi quy β_1 và β_2 cung cấp các ước lượng về mức độ ảnh hưởng của ảnh hưởng chính và β_3 là ảnh hưởng tương tác. Ngoài ra, cùng một mô hình thừa số có thể được phân tích bằng cách sử dụng một phân tích ANOVA hai chiều. Phân tích hồi quy liên quan đến nhiều biến dự báo đôi khi được gọi là hồi quy bội số, nó khác với hồi quy đa biến sử dụng nhiều biến kết quả.

Một lưu ý khi giải thích các tác động tương tác. Nếu tác động tương tác β_3 là đáng kể (có ý nghĩa) thì ảnh hưởng của thử nghiệm (phương pháp giảng dạy) đối với kết quả học tập của sinh viên phụ thuộc vào thời lượng giảng dạy. Trong trường hợp này, chúng ta không thể đưa ra những giải thích thuyết phục về ảnh hưởng độc lập của từng công cụ thử nghiệm β_1 hay β_2 bởi vì chúng không bao giờ có thể tách rời, chúng luôn tương tác với nhau để tác động (nếu có) lên biến phụ thuộc. Vì vậy, ảnh hưởng chính chỉ có thể được giải thích khi hiệu ứng tương tác là không đáng kể.

Đồng biến có thể xuất hiện trong thiết kế thừa số như là các biến mới, với hệ số hồi quy mới (ví dụ, β_4). Đồng biến có thể được đo bằng cách sử dụng các thang khoảng hoặc tỷ lệ, ngay cả khi các yếu tố dự báo được thiết kế như là các biến giả. Giải thích các đồng biến cũng theo quy tắc tương tự như của bất kỳ biến dự báo khác.

Các phân tích định lượng khác

Có rất nhiều kỹ thuật suy luận thống kê hữu ích khác dựa trên các biến thể của mô hình tuyến tính tổng quát GLM, do mục tiêu của cuốn sách này là hướng tới sự khái quát và cô đọng nên các kỹ thuật này sẽ được đề cập ngắn gọn ở đây. Bạn đọc quan tâm nên tham khảo thêm các sách chuyên khảo hoặc tham dự các khóa học thống kê để biết thêm thông tin chi tiết về các kỹ thuật này:

- *Phân tích nhân tố* (factor analysis) là kỹ thuật thu nhỏ dữ liệu, nó tổng hợp một số lượng lớn các chỉ số đã quan sát thành một tập hợp nhỏ các biến không quan sát được (tiềm tàng) dựa trên mô hình tương quan lưỡng biến cơ bản. Các biến tiềm tàng này được gọi là các nhân tố. Kỹ thuật này được sử dụng rộng rãi để đánh giá giá trị hội tụ và giá trị phân biệt trong các phương thức đánh giá (thang đo lường) đa chỉ số của nghiên cứu khoa học xã hội.

- *Phân tích biệt số* (discriminant analysis) là một kỹ thuật phân loại nhằm mục đích đặt một quan sát vào một trong các phân nhóm định danh dựa trên sự kết hợp tuyến tính của các biến dự đoán. Kỹ thuật này tương tự như kỹ thuật hồi quy đa biến, chỉ khác ở chỗ biến phụ thuộc của phân tích biệt số là định danh. Nó được sử dụng phổ biến trong các ứng dụng tiếp thị, chẳng hạn như để phân loại khách hàng hoặc các sản phẩm thành các loại khác nhau dựa trên các thuộc tính nổi bật được xác định trong cuộc điều tra quy mô lớn.

- *Hồi quy logistic* (hoặc mô hình logit) là một mô hình hồi quy tuyến tính GLM trong đó biến kết quả là nhị phân (0 hoặc 1), tuân theo một phân phối tương trưng nhằm mục tiêu dự đoán xác suất thành công của kết quả bằng cách hướng dữ liệu vào một đường cong logistic. Một ví dụ là dự đoán xác suất của các cơn đau tim trong một khoảng thời gian cụ thể, dựa trên các yếu tố dự báo như tuổi tác, chỉ số trọng lượng cơ thể, chế độ tập luyện,... Hồi quy logistic được ứng dụng rất phổ biến trong y khoa. Ước lượng về mức độ ảnh hưởng dựa trên "tỷ số vênh", biểu thị tỷ lệ chênh lệch của một sự kiện xảy ra trong một nhóm so với nhóm khác.

- *Hồi quy probit* (hoặc mô hình probit) là một GLM trong đó biến kết quả có thể thay đổi giữa 0 và 1 (hoặc có thể giả định các giá trị rời rạc 0 và 1), tuân theo một phân phối chuẩn, với mục tiêu dự đoán xác suất của mỗi kết quả đầu ra. Đây là một kỹ thuật phổ biến để phân tích dự báo trong ngành thống kê bảo hiểm, dịch vụ tài chính,... với các ứng dụng như xác định điểm tín dụng dựa trên xếp hạng tín dụng của từng người, tiền lương, nợ và các thông tin khác từ hồ sơ vay vốn của người đó. Hồi quy probit và hồi quy logit thường có hệ số hồi quy tương tự trong các ứng dụng so sánh (kết quả nhị phân) tuy nhiên mô hình logit dễ dàng hơn để tính toán và giải thích.

- *Phân tích đường dẫn* (path analysis) là một kỹ thuật GLM đa biến để phân tích mối quan hệ định hướng trong một tập hợp các biến. Nó cho phép kiểm tra các mô hình tương tác phức tạp mà ở đó biến phụ thuộc trong một phương trình này lại là biến độc lập trong phương trình khác và nó được sử dụng rộng rãi trong nghiên cứu khoa học xã hội hiện đại.

- *Phân tích chuỗi thời gian* (time series analysis) là một kỹ thuật để phân tích dữ liệu thời gian theo chuỗi thời gian hoặc các biến liên tục thay đổi theo thời gian. Ví dụ của ứng dụng này bao gồm các dự báo biến động thị trường chứng khoán và tỷ lệ tội phạm trong các đô thị. Kỹ thuật này phổ biến trong kinh tế, tài chính kế toán và xử lý tín hiệu. Kỹ thuật đặc biệt được sử dụng để hiệu chỉnh những tương quan tự động hoặc tương quan trong phạm vi giá trị của cùng một biến theo thời gian.

Chương 16

ĐẠO ĐỨC TRONG NGHIÊN CỨU KHOA HỌC

Từ điển Webster định nghĩa đạo đức là sự phù hợp với các chuẩn mực hành vi của một hoặc của một nhóm nghề nghiệp nhất định. Những chuẩn mực như vậy thường được quy định một bộ quy tắc ứng xử nghề nghiệp và nhiều khi được thực thi bởi một ủy ban tại các trường đại học - được gọi là Hội đồng Thẩm tra đạo đức (Institutional Review Board). Kể cả trong trường hợp các nguyên tắc đạo đức không được quy định rõ ràng trong các văn bản, các nhà khoa học vẫn phải có ý thức thực hiện những thỏa thuận chung của cộng đồng khoa học về những hành vi nào là chấp nhận được, hành vi vào không chấp nhận được trong tiến hành nghiên cứu khoa học chuyên nghiệp. Ví dụ, các nhà khoa học không nên thực hiện những hành vi thu thập, phân tích và xử lý dữ liệu trái với các nguyên tắc của khoa học hoặc phương pháp khoa học.

Tại sao đạo đức trong nghiên cứu khoa học lại được coi trọng như vậy? Bởi lẽ khoa học thường bị cá nhân hay tổ chức lợi dụng một cách phi đạo đức để phục vụ lợi ích riêng và tiến hành những hoạt động trái với các chuẩn mực ứng xử khoa học. Một ví dụ kinh điển là việc thử nghiệm dược phẩm Vioxx của hãng dược khổng lồ Merck khi công ty này đã che giấu cộng đồng khoa học các tác dụng phụ gây chết người của loại thuốc này. Kết quả là 3.468 người dùng Vioxx đã chết, nguyên nhân chủ yếu là do tim ngừng đập. Năm 2010, công ty đã phải đồng ý chi trả bồi thường 4,85 tỷ USD và thành lập hai ủy ban độc lập và một ban giám sát y tế để giám sát sự an toàn của quá trình nghiên cứu thuốc. Hành vi của công ty Merck là phi đạo đức và vi phạm các nguyên tắc khoa học về thu thập, phân tích và giải thích dữ liệu.

Đạo đức học là sự phân biệt giữa đúng và sai về mặt đạo đức. Những gì là phi đạo đức chưa hẳn là bất hợp pháp. Nếu hành vi một của nhà khoa học rơi vào vùng “xám” giữa đạo đức và luật pháp (nghĩa là hợp pháp nhưng phi đạo đức), thì nhà khoa học đó có thể không bị truy cứu theo quy định pháp luật, nhưng có thể bị tẩy chay trong cộng đồng khoa học, đối mặt với suy giảm nghiêm trọng về uy tín nghề nghiệp và thậm chí có thể bị sa thải vì vi phạm nguyên tắc nghề nghiệp. Những chuẩn mực đạo đức này có thể không giống nhau ở các xã hội khác nhau. Ở đây, chúng tôi đề cập đến tiêu chuẩn đạo đức áp dụng phổ biến cho nghiên cứu khoa học ở các nước phương Tây.

Các nguyên tắc đạo đức trong nghiên cứu khoa học

Dưới đây là một số những nguyên tắc về hành vi đạo đức được chấp nhận rộng rãi trong cộng đồng khoa học.

Tham gia tự nguyện và vô hại (voluntary participation and harmlessness). Những người tham gia trong một dự án nghiên cứu phải nhận thức được rằng sự tham gia của họ trong nghiên cứu này là tự nguyện; họ hoàn toàn tự do rút khỏi nghiên cứu bất cứ lúc nào mà không phải gánh chịu bất kỳ hậu quả gì; họ không thể bị tổn hại dù có tham gia hay không tham gia dự án. Vi phạm trắng trợn nhất về nguyên tắc tham gia tự nguyện có lẽ là các thí nghiệm y học cưỡng bức do các nhà nghiên cứu Đức Quốc xã thực hiện đối với các tù nhân chiến tranh trong Thế chiến II. Hành vi này đã bị Tòa án Nuremberg sau chiến tranh kết tội (những thí nghiệm này cũng là nguồn gốc của thuật ngữ “tội ác chống lại loài người”).

Các hành vi vi phạm khác ít được biết hơn như thí nghiệm bệnh giang mai tại thành phố Tuskegee do Cơ quan y tế công cộng của Mỹ tiến hành trong các năm 1932 -1972. Có gần 400 người đàn ông người Mỹ gốc Phi nghèo khổ nhiễm bệnh giang mai bị từ chối điều trị, ngay cả khi thuốc kháng sinh penicillin đã được thử nghiệm thành công điều trị hiệu quả bệnh giang mai. Những người bản cùng này bị buộc điều trị bằng các phương pháp sai lầm khác như chích mủ cột sống. Cho dù số người này không phải đối mặt với cái chết, nhưng không nên buộc họ phải chịu đựng sự đau đớn khi họ tham gia vào thí nghiệm. Năm 1971, nhà tâm lý học Philip Zimbardo tiến hành *Thí nghiệm nhà tù Stanford*. Khi đó các học sinh tại thành phố Stanford được tuyển làm đối tượng nghiên cứu và được phân công ngẫu nhiên trong vai trò tù nhân hoặc cai ngục. Nghiên cứu này sau đó đã gây nên những hậu quả tổn thương tâm lý nghiêm trọng cho các học sinh đóng vai tù nhân do hậu quả của việc bị giam giữ giả tạo. Còn những học sinh trong vai cai ngục đã thể hiện tính hung ác và ảnh hưởng đến sự nhận thức của họ về sau. Thí nghiệm đã bị buộc chấm dứt.

Ngày nay, nếu một giảng viên yêu cầu sinh viên trong lớp hoàn thành một bảng câu hỏi điều tra và thông báo rằng sự tham gia của họ là tự nguyện thì phải đảm bảo chắc chắn sinh viên không phải lo sợ nếu không tham gia có thể bị điểm thấp. Ví dụ, sẽ là phi đạo đức nếu thưởng điểm cho những sinh viên tham gia và không thưởng cho những người không tham gia, bởi vì nó làm cho người không tham gia bị bất lợi. Để tránh những trường hợp như vậy, người giáo viên có thể phải giao một nhiệm vụ khác cho những sinh viên không tham gia để họ có thể có được điểm thưởng mà không cần tham gia vào nghiên cứu

hoặc giảng viên phải thưởng điểm cho tất cả mọi người, không phân biệt họ tham gia hoặc không tham gia nghiên cứu.

Hơn nữa, trước khi xác nhận các câu trả lời tham gia, họ phải được nhận và ký vào bản **tự nguyện với đầy đủ thông tin** (informed consent). Bản này ghi rõ ràng quyền không tham gia và quyền rút khỏi nghiên cứu của họ. Trong nghiên cứu về y học, bản này cũng phải chỉ rõ rủi ro bất kỳ có thể xảy ra trong quá trình tham gia của họ. Đối với người dưới 18 tuổi, mẫu đơn này phải có chữ ký của cha mẹ hoặc người giám hộ hợp pháp. Các nhà nghiên cứu phải giữ lại các bản chấp thuận tự nguyện trong khoảng thời gian nhất định (thường là ba năm) sau khi hoàn thành quá trình thu thập dữ liệu phù hợp với các quy định về nghiên cứu khoa học trong lĩnh vực hoặc nơi làm việc của họ.

Ẩn danh và bảo mật (anonymity and confidentiality). Để bảo vệ quyền lợi hiện tại và tương lai của người tham gia, danh tính của họ phải được bảo vệ trong nghiên cứu khoa học. Điều này được thực hiện bằng cách sử dụng các nguyên tắc kếp về ẩn danh (giấu tên) và bảo mật. **Ẩn danh** đặt ra yêu cầu rằng các nhà nghiên cứu hoặc độc giả của báo cáo nghiên cứu hoặc bài báo khoa học cuối cùng không thể xác định bất kỳ câu trả lời nào là của một người trả lời cụ thể. Một ví dụ giấu tên trong nghiên cứu khoa học là một cuộc điều tra bằng thư, trong đó không có những địa chỉ giúp xác định ai trả lời và ai không trả lời. Đặc biệt, trong nghiên cứu về các hành vi lệch chuẩn, chẳng hạn như sử dụng ma túy hoặc tải nhạc bất hợp pháp của sinh viên, các câu trả lời trung thực có thể không thể đạt được nếu nhà nghiên cứu không bảo đảm tên của những người tham gia được giữ kín. Hơn nữa, giấu tên để đảm bảo rằng người tham gia được bảo vệ và tránh bị phiền hà của các cơ quan chức năng hoặc chính quyền có thể để ý tới việc nhận diện và tìm kiếm những người này trong tương lai.

Trong một số thiết kế nghiên cứu, chẳng hạn như các cuộc phỏng vấn trực diện, việc giấu tên là không thể được. Trong các thiết kế khác, chẳng hạn như khảo sát theo chiều dọc, giấu tên cũng khó được đảm bảo, bởi nếu giấu tên các nhà nghiên cứu không thể thực hiện các so sánh về phản ứng của cùng một người ở các điểm khác nhau để phân tích theo tuyến dọc. Trong những trường hợp như vậy, người tham gia cần được **bảo mật**, trong đó chỉ riêng nhà nghiên cứu có thể nhận ra phản ứng của một người nào đó, nhưng phải đảm bảo sẽ không tiết lộ danh tính người đó trong bất kỳ báo cáo, bài báo khoa học hay diễn đàn công cộng nào. Bảo mật là một hình thức bảo vệ thấp hơn so với hình thức giấu tên, bởi vì trong luật pháp Hoa Kỳ, dữ liệu thu thập trong nghiên cứu xã hội không được hưởng các “đặc quyền thông tin” như với cơ chế trao đổi thông tin của linh mục hoặc luật sư. Ví

dụ, hai năm sau khi tàu chở dầu khổng lồ Exxon Valdez tràn 10 triệu thùng dầu thô gần cảng Valdez ở bang Alaska, các cộng đồng bị thiệt hại kinh tế và môi trường ở đây đã uỷ quyền cho một tổ chức nghiên cứu ở San Diego khảo sát các hộ gia đình bị ảnh hưởng để tìm hiểu chi tiết những khó khăn và căng thẳng tâm lý gia tăng trong các gia đình này. Do các quan niệm văn hóa, nhiều người Mỹ bản địa cảm thấy rất không thoải mái nếu thông tin về thực tại cuộc sống gia đình họ bị công khai, họ được đảm bảo bí mật về danh tính khi tham gia nghiên cứu. Khi bằng chứng từ nghiên cứu được trình lên tòa án, bị đơn là công ty Exxon đã kiến nghị tòa án cần phải trưng ra các bảng câu hỏi khảo sát gốc (với thông tin nhận dạng) để kiểm tra chéo câu trả lời của những người tham gia với danh tính được bảo mật và toà đã chấp nhận yêu cầu này. May mắn thay, vụ án Exxon Valdez được giải quyết trước khi các nạn nhân (người tham gia nghiên cứu) bị buộc phải ra làm chứng tại tòa án công khai. Cho dù vậy thì khả năng vi phạm tính bảo mật tương tự như trên vẫn tiềm tàng.

Trong một trường hợp hi hữu khác, Rick Scare - một sinh viên tốt nghiệp trường Đại học Washington, đã tiến hành các nghiên cứu quan sát tham dự cùng với các nhà hoạt động bảo vệ động vật. Năm 1990, Scare xuất bản cuốn sách *Ecowarriors: Understanding the Radical Environmental Movement* (Chiến sĩ bảo vệ sinh thái: Tìm hiểu về phong trào môi trường cấp tiến). Năm 1993, Scare được gọi ra trước bồi thẩm đoàn để hỏi về các nhà hoạt động môi trường mà anh đã cộng tác nghiên cứu. Để làm tròn bổn phận đạo đức với tư cách là một thành viên của Hiệp hội xã hội học Hoa Kỳ, Scare đã chối trả lời câu hỏi bồi thẩm đoàn. Vì vậy, anh bị phạt tù 159 ngày tại Trại giam hạt Spokane. Để tự bảo vệ mình, tránh những tai nạn tương tự như Rik Scare, các nhà nghiên cứu nên xóa bỏ tất cả các thông tin cá nhân trong các tài liệu và các tệp dữ liệu ngay khi chúng không còn cần thiết. Năm 2002, Bộ Y tế và Dịch vụ Nhân sinh Hoa Kỳ đã ban hành “*Giấy chứng nhận bảo mật*” để bảo vệ các thành viên tham gia các dự án nghiên cứu, giúp họ tránh những phiền hà từ lực lượng cảnh sát và các cơ quan có thẩm quyền khác. Không phải tất cả các dự án nghiên cứu hội đủ điều kiện để đạt được sự bảo vệ này, nhưng trong nhiều trường hợp, điều này có thể cung cấp những hỗ trợ quan trọng để bảo vệ bí mật cho người tham gia.

Cung cấp thông tin về đề tài nghiên cứu (disclosure). Thông thường, trước khi thu thập dữ liệu, các nhà nghiên cứu có nghĩa vụ cung cấp một số thông tin về đề tài nghiên cứu cho những người tham gia tiềm năng để họ quyết định có tham gia vào nghiên cứu này hay không. Ví dụ, người tham gia tiềm năng cần biết những ai đang tiến hành nghiên cứu, nhằm mục đích gì, những kết quả dự kiến ra sao và những lợi ích mà kết quả nghiên cứu

mang lại. Tuy nhiên, trong một số trường hợp, việc tiết lộ thông tin như vậy có thể có khả năng mang lại những câu trả lời thiếu khách quan. Ví dụ, nếu mục đích của nghiên cứu là để xem xét khi nào một cá nhân sẽ từ bỏ quan điểm cá nhân của mình để tuân theo “quan điểm tập thể”. Theo kế hoạch, họ sẽ tham gia vào một thí nghiệm, họ lắng nghe ý kiến của người khác về một chủ đề nhất định trước khi bày tỏ của quan điểm riêng mình. Như vậy, việc tiết lộ mục đích của cuộc nghiên cứu trước khi tiến hành dẫn đến khả năng làm cho người tham gia điều chỉnh hành vi của họ trong quá trình thực nghiệm. Trong những trường hợp như vậy, mục đích của nghiên cứu chỉ nên tiết lộ trong một buổi gặp gỡ ngay sau quá trình thu thập dữ liệu, với một danh sách các khả năng rủi ro hoặc tổn hại mà người tham gia có thể gặp phải trong quá trình thực nghiệm.

Phân tích và báo cáo (*Analysis and reporting*). Các nhà nghiên cứu cũng có nghĩa vụ đạo đức đối với cộng đồng khoa học về dữ liệu được phân tích và báo cáo trong nghiên cứu của họ. Những phát hiện không mong đợi hoặc tiêu cực nên được tiết lộ đầy đủ, ngay cả khi chúng gây ra sự hoài nghi vào giá trị của thiết kế nghiên cứu hay các kết luận của đề tài nghiên cứu. Tương tự như vậy, yêu cầu này cũng áp dụng với những kết quả đáng chú ý được phát hiện sau khi nghiên cứu đã hoàn thành, cho dù những phát hiện này là do ngẫu nhiên hay do thiếu dữ liệu. Việc khẳng định những kết quả ngẫu nhiên đó là sản phẩm có chủ ý, trong kế hoạch của nhà nghiên cứu là hành vi phi đạo đức. Nói cách khác, trong nghiên cứu thực chứng, không nên hình thành các giả thuyết sau khi đã có kết quả phân tích dữ liệu, bởi vì vai trò của dữ liệu trong nghiên cứu như vậy là để kiểm tra giả thuyết, chứ không phải để xây dựng chúng.

Nó cũng là phi đạo đức khi “cắt xén” dữ liệu thành các phân đoạn khác nhau để chứng minh hay bác bỏ các giả thuyết, hay để tạo ra nhiều sản phẩm khoa học mà các sản phẩm đó đòi hỏi nguồn các dữ liệu khác nhau. Tự động nhào nặn những kết quả còn nghi vấn trở thành những kết quả nghiên cứu hợp lệ bằng cách phân tích những dữ liệu cắt xén, cục bộ, không đầy đủ hoặc không thích hợp cũng là một hình thức không trung thực trong khoa học. Tiến bộ khoa học có được thông qua sự cởi mở và trung thực, vậy nên cách tốt nhất để các nhà nghiên cứu đóng góp cho khoa học và cộng đồng khoa học là tiết lộ đầy đủ những vấn đề còn tồn tại trong nghiên cứu của họ. Qua đó, có thể giúp các nhà nghiên cứu khác tránh những vấn đề tương tự.

Hội đồng Thẩm tra đạo đức (Institutional Review Board)

Đạo đức nghiên cứu liên quan đến đối tượng là con người được quản trị bởi đạo luật

liên bang của Hoa Kỳ. Bất kỳ cơ quan nghiên cứu nào, chẳng hạn như trường đại học hoặc bệnh viện, muốn đề xuất xin ngân sách liên bang để hỗ trợ các dự án nghiên cứu phải chứng minh được rằng nghiên cứu phù hợp với pháp luật liên bang về quyền người tham gia. Quá trình này được giám sát bởi một hội đồng gồm các chuyên gia được gọi là Hội đồng Thẩm tra đạo đức (Institutional Review Board - IRB). IRB xem xét tất cả các đề xuất nghiên cứu có đối tượng tham gia là con người để đảm bảo rằng các nguyên tắc về tham gia tự nguyện, vô hại, giấu tên, bảo mật,... được thực thi, giảm thiểu những rủi ro có thể đối với người tham gia. Mặc dù pháp luật liên bang chỉ áp dụng riêng cho các đề án liên bang tài trợ, tuy nhiên các tiêu chuẩn và thủ tục này được áp dụng cho cả các đề tài không được tài trợ và cả đề tài của sinh viên.

Quá trình phê duyệt của IRB đòi hỏi nhà nghiên cứu hoàn thành một hồ sơ theo mẫu cung cấp đầy đủ thông tin về dự án nghiên cứu, các thành viên nghiên cứu (các nhà điều tra chính), chi tiết về quyền của người tham gia sẽ được bảo vệ. Những tài liệu bổ sung như bản chấp thuận tự nguyện, bảng hỏi hoặc cách thức phỏng vấn cũng có thể phải cung cấp. Các nhà nghiên cứu trong đề tài cũng phải chứng minh được rằng họ đã hiểu và nắm bắt sâu sắc các nguyên tắc đạo đức trong nghiên cứu bằng cách cung cấp giấy chứng nhận đã tham dự vào một khóa học về đạo đức nghiên cứu. Thu thập dữ liệu chỉ có thể bắt đầu sau khi dự án được IRB chấp thuận.

Bộ Quy tắc đạo đức nghề nghiệp

Hầu hết các cơ quan, tổ chức nghiên cứu chuyên nghiệp đều xây dựng và công bố các quy định chính thức về hoạt động khoa học trong đó nêu rõ những gì là được chấp nhận và không được chấp nhận trong nghiên cứu của các thành viên. Ví dụ, quy định của Hiệp hội Hệ thống thông tin (AIS), một hiệp hội nghề nghiệp toàn cầu của các nhà nghiên cứu chuyên ngành hệ thống thông tin, được tóm tắt trong Bảng 16.1 (có thể tham khảo đầy đủ quy định tại địa chỉ trực tuyến <http://home.aisnet.org/displaycommon.cfm?an=1&subarticlenbr=15>). Các quy định tương tự về đạo đức nghiên cứu cũng có thể áp dụng cho các chuyên ngành khác.

Bộ Quy tắc của AIS nhóm các hành vi vi phạm đạo đức thành hai loại. Loại I bao gồm các vi phạm nghiêm trọng như đạo văn và giả mạo dữ liệu, làm sai lệch quy trình nghiên cứu và phân tích dữ liệu. Những vi phạm này có thể dẫn đến sự trục xuất ra khỏi hiệp hội, sa thải, truy cứu pháp lý và gây tổn hại nghiêm trọng danh dự nghề nghiệp. Loại II bao gồm sự vi phạm ít nghiêm trọng hơn như không tôn trọng các quyền của đối tượng

nghiên cứu, trình bày sai lệch đóng góp của dự án nghiên cứu, sử dụng các dữ liệu của những người khác đã công bố mà không có sự chú dẫn. Những hành vi này có thể làm tổn hại cho uy tín nghề nghiệp, các công bố bị loại bỏ khỏi các tạp chí khoa học, v.v... Quy định này cũng cung cấp chỉ dẫn về những hành vi nghiên cứu chuẩn mực, cách xử lý khi phát hiện có sự vi phạm đạo đức (cho cả người vi phạm và nạn nhân) và trình tự, thủ tục giải quyết các trường hợp vi phạm đạo đức của AIS. Mặc dù những quy định về đạo đức nghiên cứu như vậy không loại trừ hoàn toàn các hành vi phi đạo đức nhưng chắc chắn, chúng giúp tạo ra ranh giới của hành vi đạo đức trong cộng đồng khoa học, giúp giảm các trường hợp vi phạm.

LOẠI MỘT: Quy định trong loại này phải **luôn luôn** được tuân thủ và nếu không sẽ cấu thành một hành vi vi phạm đạo đức nghiêm trọng. Những vi phạm nghiêm trọng này có thể dẫn đến việc trục xuất bạn ra khỏi các hiệp hội học thuật, sa thải, truy cứu pháp lý đối với bạn và có thể làm phương hại nghiêm trọng đến uy tín khoa học của bạn.

1. Không được đạo văn.

2. Không làm giả hoặc làm sai lệch dữ liệu, quy trình nghiên cứu và phân tích dữ liệu.

LOẠI HAI: Quy định trong loại này là những hành vi đạo đức được khuyến cáo. Vi phạm trắng trợn những quy định này hoặc các quy ước nghề nghiệp khác dù ít nghiêm trọng hơn, nhưng vẫn có thể dẫn đến huỷ hoại danh tiếng của bạn, bị xử lý bởi các ban biên tập, rắc rối nghề nghiệp, truy cứu pháp lý và làm giảm uy tín của bạn với đồng nghiệp.

3. Tôn trọng các quyền của đối tượng nghiên cứu, đặc biệt là quyền về thông tin riêng tư và quyền được thông báo về bản chất của nghiên cứu và các loại hoạt động họ được đề nghị tham gia.

4. Không trình bày sai lệch về đóng góp của nghiên cứu với các biên tập viên và những người tổ chức chương trình hội thảo khi bạn gửi kết quả nghiên cứu cho họ.

5. Không lạm dụng quyền hạn và trách nhiệm khi bạn được giao làm người biên tập, bình duyệt hoặc giám sát và đảm bảo rằng các quan hệ cá nhân không tác động đến nhận xét đánh giá của bạn.

6. Khai báo tất cả xung đột lợi ích có thể ảnh hưởng tới sự khách quan và vô tư của bạn khi được giao thẩm định những đề xuất công bố, các hồ sơ đăng ký thực hiện đề tài và những phần mềm hoặc khi được giao thực hiện các đề tài từ nguồn kinh phí bên ngoài.
7. Không sử dụng dữ liệu đã công bố của người khác mà không có sự trích dẫn hoặc sử dụng dữ liệu chưa công bố mà không có sự cho phép và sự thừa nhận của tác giả.
8. Ghi nhận những đóng góp có ý nghĩa của tất cả người tham gia nghiên cứu tùy theo đóng góp trí tuệ của họ, cho dù là các đồng nghiệp hay sinh viên.
9. Không sử dụng các bài viết, thông tin, ý tưởng, khái niệm hoặc dữ liệu chưa được công bố của người khác mà không có sự cho phép của tác giả khi bạn cho rằng đó có thể là kết quả của một quá trình học thuật ví dụ nhưng quá trình bình duyệt các công bố khoa học.
10. Chỉ sử dụng các tài liệu lưu trữ phù hợp với các quy định của các cơ sở lưu trữ.

LỜI KHUYÊN: Một số gợi ý về cách thức tự bảo vệ khỏi các tranh chấp quyền tác giả, sơ xuất, sai lầm và thậm chí bị truy cứu pháp lý.

1. Giữ các tài liệu và dữ liệu cần thiết nhằm xác nhận quyền tác giả của mỗi công trình học thuật khi bạn tham gia nghiên cứu.
2. Không công bố những ý tưởng khoa học cũ của chính bạn như thể chúng là một đóng góp trí tuệ mới.
3. Giải quyết các vấn đề về quyền sở hữu dữ liệu trước khi xử lý, phân tích chúng.
4. Trao đổi ý kiến các đồng nghiệp cùng chuyên ngành nếu nghi ngờ.

Bảng 16.1. Quy tắc đạo đức của Hiệp hội Hệ thống thông tin

Một tranh luận về đạo đức

Robert Allen Humphreys “Laud” là nhà xã hội học và là tác giả người Mỹ, được biết đến nhiều nhất bởi luận án tiến sĩ của ông: *Thương mại phòng trà (Tearoom Trade)*, xuất bản năm 1970. Cuốn sách này là một đề tài dân tộc học, nghiên cứu các cuộc gặp gỡ chớp nhoáng giữa những nam đồng tính ẩn danh trong nhà vệ sinh công cộng trong công viên - một hiện tượng được gọi là “phòng trà” trong tiếng lóng của dân “gay” ở Hoa Kỳ.

Humphreys bị hấp dẫn bởi một thực tế là phần lớn những người tham gia vào các hoạt động “phòng trà” về bề ngoài rất nam tính và là những người đàn ông sống một cuộc sống gia đình bình thường trong cộng đồng của họ. Tuy nhiên, điều quan trọng là họ che giấu danh tính trong các chuyến thăm tới “phòng trà”.

Thông thường, những cuộc gặp trong “phòng trà” có ba người - hai người đàn ông tiến hành quan hệ tình dục và một người canh chừng, được gọi là “Nữ hoàng cảnh giới” (Watchqueen). Nhiệm vụ của “Nữ hoàng cảnh giới” là để cảnh báo cho hai người kia về sự xuất hiện của cảnh sát hoặc những người khác, đồng thời người này cũng nảy sinh những khoái cảm khi xem hành vi quan hệ tình dục. Để tiếp cận các đối tượng này, Humphreys xuất hiện tại các nhà vệ sinh công cộng và giả mạo như một watchqueen. Với tư cách là người quan sát tham dự, Humphreys có thể tiến hành quan sát cho phạm vi nghiên cứu luận án của mình. Ông coi việc thực hiện phương pháp quan sát tham dự này là bình thường, giống như khi nghiên cứu những hiện tượng xã hội khác chẳng hạn như hoạt động biểu tình.

Humphreys cần thêm thông tin về những người tham gia. Nhưng vì những người tham gia không muốn trả lời phỏng vấn về lĩnh vực này hoặc không muốn tiết lộ danh tính cá nhân, Humphreys đã ghi lại biển số xe ô tô của người tham gia mỗi khi có thể và lần ra tên cùng địa chỉ của họ từ cơ sở dữ liệu chung. Sau đó, ông đến thăm những người đàn ông này tại nhà của họ. Ông ngụy trang để tránh bị nhận diện và thông báo rằng ông đang tiến hành một cuộc khảo sát và sau đó ông thu thập dữ liệu cá nhân, những dữ liệu này sẽ không thể có được nếu ông không ngụy trang như vậy.

Nghiên cứu của Humphreys đã tạo ra tranh cãi khá gay gắt trong cộng đồng khoa học. Nhiều nhà phê bình nói rằng ông không nên xâm phạm quyền riêng tư của người khác khi nghiên cứu khoa học; những người khác thì quan ngại rằng Humphrey đã gian dối với những người tham gia nghiên cứu khi làm cho họ tin rằng ông chỉ là một “Nữ hoàng cảnh giới”, trong khi rõ ràng ông có những động cơ bí mật khác. Kể cả những người cho rằng hành vi quan sát trong phòng trà có thể chấp nhận được bởi vì những người tham gia đã sử dụng các công trình công cộng thì họ cũng cho rằng các cuộc phỏng vấn những người tiếp sau đó của Humphrey lại là phi đạo đức vì ông đã ngụy trang lý do gặp gỡ, vì cách ông tìm được địa chỉ nhà của họ và đã không tìm kiếm sự chấp thuận tự nguyện.

Trong khi đó, một vài nhà nghiên cứu đã bảo vệ phương pháp tiếp cận của Humphrey, họ nói rằng đây là một hiện tượng xã hội học quan trọng, xứng đáng được

nghiên cứu kỹ lưỡng; rằng không có cách nào khác để có thể thu thập được những dữ liệu có giá trị như vậy; và hơn nữa, việc nguy trang của Humphreys là vô hại, ông không tiết lộ danh tính của các đối tượng nghiên cứu của mình cho bất cứ ai. Tranh cãi này không bao giờ được phán xử thuyết phục và nó vẫn tiếp tục còn tranh luận sôi nổi trong các lớp học và các diễn đàn về đạo đức nghiên cứu./.

PHỤ LỤC

Mẫu chương trình môn học đào tạo tiến sĩ

QMB 7565: Giới thiệu về phương pháp nghiên cứu

Đại học South Florida

Trường Kinh doanh

Học kỳ mùa thu 2011

Mục tiêu:

Môn học này sẽ trang bị cho nghiên cứu sinh những kiến thức cơ bản về quá trình tiến hành nghiên cứu học thuật. Chúng ta sẽ tìm hiểu về cách tư duy và hành động của một nhà nghiên cứu khi tiến hành các bước khái niệm hóa, thiết kế, thực hiện và đánh giá các dự án nghiên cứu “khoa học”. Một phần của môn học sẽ yêu cầu bạn thiết kế và viết một đề cương nghiên cứu độc lập (với sự trợ giúp của giáo sư). Ngoài ra, các kỳ thi cuối môn học sẽ giúp bạn có được những chuẩn bị về phương pháp luận tốt nhất cho bước bảo vệ luận án tiến sĩ của bạn.

Cấu trúc:

Môn học được thiết kế theo hình thức seminar. Thảo luận và phân tích các bài đọc đã giao là tâm điểm của tất cả các buổi seminar trong khóa đào tạo nghiên cứu sinh. Muốn vậy, bạn phải đọc *tất cả* các bài đọc mà giảng viên giao cho trước khi tới lớp học, phân tích các nội dung trong bài đọc, tranh luận các nội dung này với bạn cùng lớp, sử dụng tư duy tổng hợp kết nối chúng với nhau để qua đó phát triển tư duy của một nhà nghiên cứu. Chú ý rằng nếu bạn không chuẩn bị đầy đủ cho các buổi học, chắc chắn bạn sẽ hoàn toàn mất phương hướng và sẽ không học hỏi được bất kỳ điều gì từ môn học.

Được thiết kế dành riêng cho nghiên cứu sinh, môn học này bao gồm: (1) một khối lượng công việc nặng hơn rất nhiều so với bất kỳ một khóa học thạc sĩ nào bạn đã trải qua, (2) một số lượng rất nhiều các bài đọc khô khan, và (3) một mức độ cao về tư duy phê phán và thường gây ra cảm giác chán nản. Đây không phải là một môn học dễ và bạn cũng không dễ để đạt được điểm cao trong môn này. Nếu bạn thấy yêu cầu của môn học là quá cao, bạn nên rút tên khỏi môn học này và có nghĩa là rút tên khỏi khóa học tiến sĩ.

Sách và tài liệu học tập:

Bhattacharjee, A. Social Science Research: Principles, Methods, and Practice, Ver 2.0, 2011, Free download from Blackboard (my.usf.edu). [Phiên bản 2.0, 2011, Tải miễn phí từ Tài khoản điện tử (my.usf.edu)].

Kuhn, T. J. The Structure of Scientific Revolutions, University of Chicago Press, Chicago, 1996. Bạn phải đọc toàn bộ cuốn sách này trước Tuần 4. Bản bìa mềm có giá trên Amazon chỉ là 7.5 đô la, nhìn có vẻ rẻ một cách khó hiểu, nhưng cuốn sách thật sự là một bài đọc rất học búa, vì vậy bạn nên đọc ngay từ lúc này.

Tài liệu khác: Tải về từ Tài khoản điện tử Blackboard (my.usf.edu). Tài liệu cụ thể được đề cập trong Chương trình học phía dưới.

Đánh giá kết quả môn học:

Các thành phần đánh giá kết quả môn học:

Thi cuối kỳ	35 điểm
Đề cương nghiên cứu	35 điểm
Thẩm định bài báo khoa học	10 điểm
Điểm chuyên cần	20 điểm

Mức điểm tổng: A+: 97-100; A: 92-97; A-: 90-92; B+: 88-90; B: 82-88; B-: 80-82; C+: 78-80; C: 70-78.

Thi cuối kỳ:

Đây là một buổi thi tổng hợp, diễn ra trong 3 giờ đồng hồ, trong đó bạn sẽ viết một bài luận để trả lời 3 - 4 câu hỏi tổng hợp, tương tự như khi bạn tham gia vào buổi thi tốt nghiệp khóa nghiên cứu sinh. Bạn có thể tải các câu hỏi mẫu cho buổi thi tại Tài khoản điện tử Blackboard. Bạn có thể tùy chọn viết tay hoặc đánh máy câu trả lời của bạn. Câu trả lời của bạn cần thể hiện là một (a) bản tổng hợp các tài liệu học tập được thảo luận trong suốt môn học và (b) bài phân tích biện chứng các tài liệu này, nó không đơn thuần là bản copy của những tài liệu đó hay những nhận xét của giáo sư.

Bạn sẽ được đánh giá bởi chất lượng của các câu trả lời chứ không phải là bạn đã viết bao nhiêu trang cho các câu trả lời đó. Bạn sẽ bị mất điểm nếu như câu trả lời của bạn thiếu trọng tâm (lan man, dài dòng), thiếu chi tiết cần thiết (quá chung chung hoặc không rõ nghĩa) hoặc chỉ thể hiện được những hiểu biết sơ sài về những nội dung đã thảo luận

trên lớp học. Đây là buổi thi mở, bạn có thể sử dụng tài liệu theo ý thích, nhưng nên nhớ rằng dù có mang rất nhiều tài liệu vào phòng thi nhưng chúng sẽ không có ý nghĩa gì nếu bạn không có một quá trình chuẩn bị tích cực trước buổi thi. Do số lượng khổng lồ các chủ đề và tài liệu được thảo luận trong suốt môn học, cố gắng nhồi nhét kiến thức chỉ trong một tuần trước khi thi chắc chắn là quá ít và quá muộn.

Viết đề cương nghiên cứu:

Bạn sẽ phải áp dụng kiến thức đã học về tiến trình nghiên cứu khoa học để viết một bản đề cương nghiên cứu và cuối học kỳ. Đề cương này dài 10 – 12 trang, giãn dòng đơn, không tính danh sách tài liệu tham khảo và phụ lục. Bạn có thể lựa chọn bất kỳ vấn đề nghiên cứu nào nhưng phương pháp luận thực chứng truyền thống được ưu tiên hơn. Vấn đề nghiên cứu phải mới (bạn chưa nghiên cứu trong các môn học khác hoặc chưa nghiên cứu với các giáo sư khác), phải tìm hiểu một vấn đề nghiên cứu có thật (không phải nội dung lý thuyết hay giả tưởng) và phải đạt đến chất lượng ít nhất tương đương với các tham luận trong các hội thảo khoa học. Đề cương của bạn phải gồm năm phần: (1) vấn đề nghiên cứu và tầm quan trọng của nó, (2) khảo cứu tài liệu, (3) lý thuyết và giả thuyết, (4) phương pháp nghiên cứu, và (5) kế hoạch nghiên cứu. Thu thập hoặc phân tích dữ liệu là không bắt buộc. Nhưng nếu bạn chọn phương pháp luận diễn giải thì một phần về thu thập và phân tích dữ liệu là cần thiết (và cấu trúc nêu trên sẽ thay đổi – trao đổi với tôi về những thay đổi này).

Đề cương nghiên cứu sẽ được thực hiện trong quá trình học suốt học kỳ khi chúng ta đề cập các chủ đề tương ứng trong lớp học. Điều này giúp tôi nhận xét và chỉnh sửa các lỗi của đề cương trước khi môn học kết thúc. Chú ý rằng, các bản thảo đề cương không được tính điểm, chỉ bản cuối cùng mới được tính. Bạn cũng sẽ phải bảo vệ đề cương trong khoảng 15 – 20 phút vào tuần học cuối, sau đó sẽ có 5 phút thảo luận với hình thức giống như các buổi hội thảo khoa học.

Thẩm định bài báo khoa học:

Một yêu cầu quan trọng đối với những người theo sự nghiệp học thuật là khả năng đánh giá các sản phẩm khoa học của người khác. Trong nửa cuối của kỳ học, bạn sẽ phải viết bản thẩm định một bài báo được gửi tới một tạp chí khoa học kinh doanh hàng đầu để công bố. Để hỗ trợ bạn thực hiện nhiệm vụ này, khi lên lớp, tôi sẽ giới thiệu với các bạn cách thức, quy trình viết bản thẩm định bài báo khoa học và sẽ giao cho mỗi người luyện tập, viết thử một bản thẩm định cho một bài báo gửi công bố thật sự. Sau khi các bạn viết

xong, tôi sẽ đưa cho các bạn những bản thẩm định chính thức của thẩm định viên giấu tên về những bài báo đó, như vậy bạn có thể so sánh bản thẩm định của bạn và của thẩm định viên chuyên nghiệp, từ đó thấy được mình thiếu điểm nào. Bản thẩm định được chấm điểm là bản bạn viết ở nhà và phải nộp một tuần trước buổi kiểm tra cuối kỳ. Bạn có thể phải đọc thêm một số tài liệu để nắm bắt được kiến thức cơ bản liên quan đến chủ đề thẩm định, nhưng chú ý là bạn không được phép làm cùng hay thảo luận về bài biết với bất kỳ ai cả trong và ngoài lớp học.

Việc tham dự lớp học (điểm chuyên cần):

Mỗi một buổi thảo luận sẽ phân công một “người điều hành”, người này có trách nhiệm (1) chuẩn bị một bản tóm tắt (theo mẫu) dài một trang cho các bài đọc sẽ thảo luận và (2) điều hành buổi thảo luận. Bạn sẽ nhận số điểm chuyên cần cao nhất nếu bạn gửi các bản tóm tắt của tất cả các bài đọc được giao đúng kế hoạch và tham gia nhiệt tình vào các buổi thảo luận. Làm ơn đọc một số bản tóm tắt mẫu để biết về cấu trúc của một bản tóm tắt. Nếu bạn viết những bản tóm tắt này cẩn thận, chúng sẽ rất hữu ích cho kỳ thi cuối kỳ. Photocopy bản tóm tắt của bạn và gửi cho mỗi bạn trong lớp một bản. Tôi sẽ nhận xét các bản tóm tắt này ngay trên lớp học, vì vậy bạn có thể cải thiện chúng cho những buổi thảo luận tiếp theo. Nếu bạn không thể tham gia một buổi nào đó hoặc muốn đọc một bài đọc khác, bạn có thể đổi bài đọc của mình cho một bạn trong lớp. Tuy nhiên vì không hiểu ý nhau hoặc vì một lý do nào khác mà khi đến lớp, một bài đọc không có bản tóm tắt thì sinh viên được tôi giao bài đọc đó sẽ bị trừ điểm chuyên cần. Mặc dù đã có các bản tóm tắt, mỗi sinh viên cần phải dự *tất cả* các buổi thảo luận, đọc *tất cả* tài liệu được giao và đóng góp vào *tất cả* các phần thảo luận.

Quy định của lớp học:

Việc tham dự: Tôi không chính thức tham gia tất cả các buổi thảo luận nhưng tôi sẽ có thông tin về ai đến lớp và ai không. Nếu bạn nghĩ rằng bạn không đến lớp nhiều hơn một tuần trong học kỳ này, bạn nên rút tên khỏi lớp học này.

Trung thực trong học tập: Đạo văn dưới bất kỳ hình thức nào đều không được chấp nhận và sẽ bị điểm FF [Force Fail] ngay tức khắc. Hãy đọc kỹ quy định về tính trung thực trong học thuật của University of South Florida để biết thêm thông tin về nội dung này.

Sinh viên khuyết tật: Những sinh viên yêu cầu hỗ trợ về khuyết tật nên thông báo cho tôi trong vòng hai tuần đầu của lớp học cùng với một lá thư từ Văn phòng Hỗ trợ sinh viên khuyết tật.

Điện thoại di động: Tắt nguồn điện thoại di động trong giờ học, tuy nhiên bạn có thể sử dụng trong giờ giải lao.

Chương trình môn học

Tuần 1: Giới thiệu về nghiên cứu

Nội dung môn học và giới thiệu.

Chương 1.

Tiến trình nghiên cứu và công bố khoa học.

Tuần 2: Tư duy của nhà nghiên cứu

Chương 2.

Một câu chuyện về hai bài báo khoa học:

- Fichman, R.G. and Kemerer, C.F., "The Illusory Diffusion of Innovation: An Examination of Assimilation Gaps," *Information Systems Research* (10:3), September 1999, pp. 255-275.

- Williams, L.; Kessler, R.R.; Cunningham, W.; and Jeffries, R., "Strengthening the Case for Pair Programming," *IEEE Software*, July/August 2000, pp. 19-25.

Tư duy phản biện:

- Dialog on Leadership, "Awareness is the First Critical Thing," A Conversation with Wanda Orlikowski, 1999.

Đạo đức nghiên cứu:

- AIS Code of Conduct: <http://home.aisnet.org/displaycommon.cfm?an=1&subarticlenbr=13>

- IRB Process: http://www.research.usf.edu/cs/irb_forms.htm

Tuần 3: Khảo cứu tài liệu và phân tích vĩ mô

Khảo cứu tài liệu:

- Fichman, R.G., "Information Technology Diffusion: A Review of Empirical Research," *Proceedings of the Thirteenth International Conference on Information Systems*, Dallas, 1992, 195-206.

- Alavi, M. and Leidner, D.E., "Knowledge Management and Knowledge Management Systems," *MIS Quarterly* (25:2), March 2001, pp. 107-136.

Phân tích vĩ mô:

- King, W.R, and He, J., "Understanding the Role and Methods of Meta-Analysis in IS Research," Communications of the AIS (16), 2005, pp. 665.686.

- Henard, D.H. and Szymanski, D.M., "Why Some New Products are More Successful Than Others," Journal of Marketing Research (38), August 2001, pp. 362-375.

Nhiệm vụ: hoàn thành Đề cương nghiên cứu phân: Vấn đề nghiên cứu và tầm quan trọng của nó.

Tuần 4: Khoa học luận

Chương 3.

Các mô thức trong nghiên cứu khoa học

- Kuhn, T., The Structure of Scientific Revolutions, University of Chicago Press, 1996 (Đọc toàn bộ sách).

[Chỉ tham khảo chú thích của Tiến sĩ Pajares khi bạn hoàn toàn không hiểu].

Các mô thức của khoa học xã hội:

- Krugman, P., "How Did Economists Get It So Wrong," New York Times, Sept 6, 2009.

- Gioia, D.A. and Pitre, E., "Multiparadigm Perspectives on Theory Building," Academy of Management Review (15:4), 1990, pp. 584-602.

Tuần 5: Lý thuyết trong nghiên cứu về tổ chức

Chương 4.

Vai trò của lý thuyết:

- Steinfield, C.W. and Fulk, J., "The Theory Imperative," in Organizations and Communications Technology, Janet Fulk and Charles W. Steinfield (eds.), Sage Publications, Newbury Park, CA, 1990.

Lý thuyết Đại diện (Agency theory):

- Eisenhardt, K.M., "Agency Theory: An Assessment and Review," Academy of Management Review (14:1), 1989, pp. 57-74.

Lý thuyết Chi phí giao dịch (Transaction cost theory):

- Williamson, O.E., "The Economics of Organization: The Transaction Cost

Approach," American Journal of Sociology (87:3), 1981, pp. 548-577.

Tuần 6: Lý thuyết về tổ chức (tiếp)

Lý thuyết Dựa vào nguồn lực (Resource-based theory):

- Barney, J.B., "Firm Resources and Sustained Competitive Advantage," Journal of Management (17:1), 1991, pp. 99-120.
- Priem, R.L. and Butler, J.E., "Is the Resource-Based 'View' a Useful Perspective for Strategic Management Research?," Academy of Management Review (26:1), 2001, pp. 22-40.

Lý thuyết Năng lực vận động (Dynamic capability theory):

- Teece, D.J.; Pisano, G.; and Shuen, A., "Dynamic Capabilities and Strategic Management," Strategic Management Journal (18:7), 1997, 509-533.

Nhiệm vụ: hoàn thành Đề cương nghiên cứu phân: Khảo cứu tài liệu (cùng với các phần trước, chỉnh sửa chúng khi cần thiết).

Tuần 7: Đo lường và giá trị khoa học

Chương 6 và 7.

Giá trị của thang đánh giá và giá trị đơn:

- Straub, D.W., "Validating Instruments in MIS Research," MIS Quarterly (13:2), June 1989, pp. 146-169.
- MacKenzie, S. B., Podsakoff, P. M., and Podsakoff, N. P., "Construct Measurement and Validation Procedures in MIS and Behavioral Research: Integrating New and Existing Techniques," MIS Quarterly (35:2), 2011, pp. 293-334.

Nhiệm vụ: hoàn thành Đề cương nghiên cứu phân: Lý thuyết và giả thuyết (cùng với các phần trước, chỉnh sửa chúng khi cần thiết).

Tuần 8: Nghiên cứu khảo sát

Chương 5, 8 và 9.

Ví dụ về nghiên cứu khảo sát thực địa:

- Tsai, W., "Knowledge Transfer in Intraorganizational Networks: Effects of Network Position and Absorptive Capacity on Business Unit Innovation and Performance," Academy of Management Review, 2001.

Thiên vị trong nghiên cứu khảo sát:

- Malhotra, N. K., Kim, S. S., and Patil, A., "Common Method Variance in IS Research: A Comparison of Alternative Approaches and a Reanalysis of Past Research," *Management Science* (52:12), 2006, pp. 1865-1883.

Tuần 9: Nghiên cứu thực nghiệm và ngụy thực nghiệm

Chương 10.

Ví dụ về nghiên cứu thực nghiệm thực địa:

- Hunton, J.E. and McEwen, R.A., "An Assessment of the Relation Between Analysts' Earnings Forecast Accuracy, Motivational Incentives, and Cognitive Information Search Strategy," *The Accounting Review* (72:4), October 1997, pp. 497-515.

Những vấn đề của nghiên cứu thực nghiệm:

- Jarvenpaa, S. L.; Dickson, G. W.; and DeSanctis, G., "Methodological Issues in Experimental IS Research: Experiences and Recommendations," *MIS Quarterly*, June 1985, pp. 141-156.

Tuần 10: Thẩm định kết quả nghiên cứu

Viết bản thẩm định bài báo khoa học:

- Lee, A.S., "Reviewing a Manuscript for Publication," *Journal of Operations Management* (13:1), July 1995, pp.87-92.

- Agarwal, R; Echambadi, R; Franco, A.M.; and Sarkar, M.B., "Reap Rewards: Maximizing Benefits from Reviewer Comments," *Academy of Management Journal* (49:2), 2006, pp.191-196.

Tại sao bài báo khoa học bị từ chối:

- Daft, R.L., "Why I Recommended that Your Manuscript be Rejected and What You Can Do About It," in L.L. Cummings & P.J. Frost (eds.), *Publishing in the Organizational Sciences*, 2nd ed., 1995, pp. 164-182.

Viết bản thẩm định chuẩn tắc:

- Dựa trên chỉ dẫn từ các tài liệu ở trên, viết một bản thẩm định, đánh giá một bài báo khoa học theo các quy định chuẩn. Sử dụng mẫu đánh giá [[review template](#)] cho cấu trúc bản thẩm định của bạn: Anonymous, Paper submitted to *MIS Quarterly*, 2004..

Các bản đánh giá thật của ba thẩm định viên AE và SE sẽ được gửi bằng email cho bạn sau khi bạn đã hoàn thành bản thẩm định.

Thẩm định bài báo khoa học:

- “Bài báo khoa học bí ẩn” [[mystery paper](#)] dành cho bạn để thẩm định và sẽ bàn luận ở tuần 13. Bản thẩm định này được chấm điểm, vì vậy, bạn cần lên kế hoạch và dành đầy đủ thời gian, tâm trí để hoàn thành nhiệm vụ này.

Tuần 11: Nghiên cứu trường hợp

Chương 11.

Tiến hành nghiên cứu trường hợp:

- Benbasat, I.; Goldstein, D.K.; and Mead, M., "The Case Research Strategy in Studies of Information Systems," MIS Quarterly, September 1987, pp. 369-386.

Ví dụ về nghiên cứu trường hợp:

- Beaudry, A. and Pinsonneault, A., "Understanding User Responses to Information Technology: A Coping Model of User Adaptation," MIS Quarterly (29:3), September 2005, pp. 493-524.

- Eisenhardt, K.M., "Making Fast Strategic Decisions In High-Velocity Environments," Academy of Management Journal (32:3), 1989, pp. 543-577.

So sánh phân tích thực chứng và diễn giải:

- Trauth, E.M. and Jessup, L.M., "Understanding Computer-Mediated Discussions: Positivist and Interpretive Analyses of Group Support System Use," MIS Quarterly (24:1), March 2000, pp. 43-79.

Nhiệm vụ: hoàn thành Đề cương nghiên cứu phần: Phương pháp nghiên cứu (cùng với toàn bộ các phần trước).

Tuần 12: Nghiên cứu diễn giải

Chương 12 và 13

Bản mẫu: Phân tích nội dung bằng cách sử dụng phương pháp Lý thuyết nền (Grounded Theory).

Nghiên cứu định tính:

- Shah, S.K. and Corley, K.G., "Building Better Theory by Bridging the Quantitative-Qualitative Divide," Journal of Management Studies (48:3), December 2006, pp. 1821-1835.

Nghiên cứu hành động:

- Kohli, R. and Kettinger, W., "Informating the Clan: Controlling Physician Costs and Outcomes," MIS Quarterly (28:3), September 2004, pp. 1-32.

Nghiên cứu dân tộc học:

- Barley, S.R., "Technicians in the Workplace: Ethnographic Evidence for Bringing Work into Organization Studies," Administrative Science Quarterly (41), 1996, pp. 404-411.

Tuần 13: Các phương pháp hỗn hợp

Chương 16.

Bản mẫu:

- Phân tích thống kê sử dụng phần mềm SPSS.

Phân tích dữ liệu thứ cấp:

- Chaney, P.K. and Philipich, K.L., "Shredded Reputation: The Cost of Audit Failure," Journal of Accounting Research (40:4), September 2002, pp. 1221-1245.

Mô hình hóa phân tích:

- Bayus, B.L., Jain, S., and Rao, A.G, "Truth or Consequences: An Analysis of Vaporware and New Product Announcements," Journal of Marketing Research (38), February 2001, pp. 3-13.

Nhiệm vụ: hoàn thành bản thảo định một bài báo khoa học

Tuần 14: Thi cuối học kỳ

Buổi thi mở tổng hợp cuối học kỳ, diễn ra trong lớp học với thời gian ba giờ đồng hồ. Tham khảo các câu hỏi thi mẫu [[Sample Exam Questions](#)].

Tuần 15: Sinh viên thuyết trình

Thuyết trình 20 phút cộng thêm 5 phút hỏi (Q) và trả lời (A)

Nhiệm vụ: hoàn thành Đề cương nghiên cứu bản cuối.

NGHIÊN CỨU KHOA HỌC XÃ HỘI:

NGUYÊN TẮC, PHƯƠNG PHÁP VÀ THỰC HÀNH

GS. Anol Bhattacharjee

Đại học South Florida, Tampa, Florida, USA

abhattach@usf.edu

TS. Phan Việt Phong

Học viện An ninh nhân dân, Hà Nội, Việt Nam

phongquanlykhoahoc500@gmail.com

TS. Cao Ngọc Anh

Học viện An ninh nhân dân, Hà Nội, Việt Nam

caoanhc500@gmail.com

Thông tin về sách

Cuốn sách này nhằm trang bị cho nghiên cứu sinh và học viên cao học kiến thức về tiến trình thực hiện nghiên cứu khoa học trong khoa học xã hội, kinh doanh, giáo dục học, y tế công và các ngành khác có liên quan. Cuốn sách là một nguồn tài liệu trọn vẹn, xúc tích và toàn diện, giới thiệu những khái niệm cơ bản trong nghiên cứu về hành vi, nó có thể là một nguồn tham khảo độc lập hoặc là nguồn bổ sung cho các tài liệu phục vụ các buổi thảo luận hay các môn học về phương pháp nghiên cứu cho các chương trình sau đại học. Hiện tại, cuốn sách này đang được sử dụng như là một giáo trình về phương pháp nghiên cứu ở các trường đại học trên cả sáu lục địa và sẽ sớm được dịch ra chín ngôn ngữ khác nhau.

Thông tin về tác giả

Anol Bhattacharjee là giáo sư ngành Hệ thống thông tin quản lý và là thành viên nhóm nghiên cứu Citigroup/Hidden River tại trường Đại học South Florida, Mỹ. Ông là một trong số mười nhà nghiên cứu nổi bật nhất trong ngành hệ thống thông tin trên thế giới. Trong hơn một thập kỷ qua (2001 - 2010), ông được xếp hạng thứ tám về số lượng các nghiên cứu được công bố trên hai tạp chí uy tín nhất của ngành hệ thống thông tin là MIS Quarterly và Information Systems Research. Trong 15 năm sự nghiệp nghiên cứu đã qua, tiến sĩ Bhattacharjee đã công bố trên 50 bài báo khoa học và hai cuốn sách với trên 4.000 trích dẫn trên Google Scholar. Ông là thành viên Ban biên tập MIS Quarterly trong bốn năm qua và thường xuyên được mời thuyết trình về kết quả nghiên cứu của ông hoặc được mời xây dựng các chương trình nghiên cứu mới tại nhiều trường đại học trên toàn cầu. Chi tiết về tiến sĩ Bhattacharjee được đăng tải trên trang web của ông <http://ab2020.weebly.com>.