

## Chương 2

# Phương pháp điều tra

### 2.1. Dàn mẫu và chọn mẫu

SAVY được thực hiện trên cơ sở chọn mẫu đại diện cho toàn bộ vị thành niên và thanh niên từ 14-25 tuổi sống trong hộ gia đình trên toàn quốc, theo 8 vùng kinh tế, khu vực thành thị/nông thôn. Đây là mẫu hệ thống được lựa chọn từ dàn mẫu 45.000 hộ của cuộc khảo sát mức sống hộ gia đình năm 2002 (sau đây được gọi là VLSS 2002). Mẫu SAVY là mẫu ngẫu nhiên phân tầng, được chọn nhiều bước và chọn theo cụm (cluster) đủ để đại diện ở cấp độ quốc gia cũng như khu vực thành thị, nông thôn. Xác suất chọn mẫu khu vực thành thị được tăng lên theo tỷ lệ thực tế của dân số, đặc biệt là Hà Nội và Hồ Chí Minh là hai trung tâm đô thị lớn của Việt Nam. Điều này nhằm tăng độ tin cậy cho các ước lượng thống kê cho khu vực thành thị và riêng hai thành phố Hà Nội và Hồ Chí Minh.

Dàn mẫu của cuộc điều tra SAVY bao gồm 42 tỉnh thuộc 61 tỉnh và thành phố<sup>1</sup> trong mẫu của VLSS 2002. Dàn mẫu SAVY, được chọn theo xác suất tỷ lệ với qui mô dân số (PPS) để bảo đảm tính đại diện. Tại mỗi tỉnh được chọn, tiến hành chọn các địa bàn điều tra (area). Đối tượng điều tra là tất cả các thanh thiếu niên độ tuổi 14-25 (sinh từ năm 1978 đến 1989) của 20 hộ điều tra thuộc địa bàn được chọn của VLSS 2002 bất kể đã lập gia đình hay chưa. Mẫu này đại diện cho thanh thiếu niên trong hộ gia đình nhưng không bao gồm những đối tượng thanh thiếu niên hiện đang tham gia quân đội, sống trong các trại giam, các trung tâm bảo trợ xã hội, ký túc xá sinh viên, thanh niên di cư...

Trong dàn mẫu 61 tỉnh thành của VLSS có 2.250 địa bàn, và trong mẫu 42 tỉnh của điều tra SAVY có 1.643 địa bàn đã tiến hành khảo sát mức sống dân cư năm 2002. Có 446 địa bàn trong số đó được chọn tham gia điều tra SAVY. Tại 446 địa bàn này 8.920 hộ tương đương với dân số 40.140 (khoảng 4,5 người trên 1 hộ). Vì thanh thiếu niên tuổi 14-25 chiếm 24,5% tổng dân số (theo số liệu của điều tra dân số 1999) ước tính số thanh thiếu niên trong mẫu của SAVY là 9.835. Nếu tỷ lệ huy động (phần trăm thực sự được phỏng vấn) đạt 90% thì số phỏng vấn là 8.850. Thực tế tỷ lệ huy động của SAVY đạt 85% và

số thanh thiếu niên thực tế được phỏng vấn là 7.584 người.

Mẫu được chọn điều tra là mẫu đại diện và quy mô đủ để phân tích ở cấp quốc gia, phân tích theo khu vực thành thị-nông thôn, giới tính và 8 vùng.

Tỷ lệ huy động tham gia điều tra SAVY là không thấp theo tiêu chuẩn của các cuộc điều tra thanh niên, và chỉ tương đối thấp hơn so với dự kiến của Tổng cục Thống kê. Lý do làm giảm tỷ lệ huy động tham gia trong cuộc điều tra SAVY chủ yếu do biến động tương đối lớn về nơi cư trú do thanh thiếu niên di chuyển từ vùng này sang vùng khác, dàn mẫu được chọn ra từ cuộc điều tra khác và sử dụng danh sách hộ gia đình lập từ một năm trước, cách thức Ủy ban nhân dân xã phường vận động thanh thiếu niên đến địa điểm phỏng vấn có thể còn làm cho thanh thiếu niên e ngại hoặc chưa thoải mái tham gia điều tra. Đây là cuộc điều tra đầu tiên áp dụng phương pháp mới: phỏng vấn thanh thiếu niên tại một địa điểm công cộng tại địa phương, kết hợp phỏng vấn trực tiếp và tự điền, trong quá trình đó không có cha mẹ hoặc người lớn tham dự. Đối với nhóm thanh thiếu niên đồng ý tham gia điều tra, tỷ lệ từ chối trả lời các câu hỏi nhạy cảm trong phần tự điền rất thấp. Phương pháp điều tra (kết hợp phỏng vấn và tự điền), chất lượng điều tra viên và tổ chức chỉ đạo điều tra, giám sát điều tra thực địa là những yếu tố quan trọng đã góp phần nâng cao chất lượng dữ liệu SAVY.

### 2.2. Phiếu hỏi

Phiếu hỏi được thiết kế rất kỹ lưỡng dựa vào kinh nghiệm từ các cuộc điều tra tương tự ở các nước trong khu vực cũng như ý kiến đóng góp của thanh thiếu niên để đảm bảo tính phù hợp và chất lượng của câu hỏi. Quá trình thiết kế phiếu hỏi cũng giúp xác định phương pháp và gợi ý cho việc triển khai công tác điều tra thực địa.

Phiếu hỏi được biên soạn có sự tham gia rộng rãi của các ban ngành, tổ chức trong nước và quốc tế, nhằm đảm bảo kết quả điều tra là sản phẩm chung của tất cả các bên hữu quan.

Nhìn chung, quá trình biên soạn phiếu hỏi chia làm 2 giai đoạn:

**Giai đoạn 1:** các nhà nghiên cứu có kinh nghiệm

trong và ngoài nước, các tổ chức hữu quan tham gia hội thảo thiết kế phiếu hỏi do Bộ Y tế tổ chức. Các đại biểu bàn bạc các chủ đề cần quan tâm cũng như cách đặt câu hỏi, đồng thời tham khảo ngân hàng câu hỏi từ các điều tra tương tự trong khu vực. Vì có một số chủ đề khá nhạy cảm, nên các đại biểu đề nghị nên chia phiếu hỏi ra làm 2 phần: 1 phần để phỏng vấn trực tiếp, 1 phần để người được phỏng vấn tự điền. Trên cơ sở đó, bản thảo phiếu hỏi đã được thiết kế để lấy ý kiến đóng góp của các đại biểu trong hội thảo, các cơ quan tham gia và đồng thời của thanh thiếu niên thông qua các cuộc tham khảo ý kiến.

8 cuộc thảo luận nhóm được tổ chức ở Hà Nội, thành phố Hồ Chí Minh với sự tham gia của 60 thanh thiếu niên trong và ngoài trường học tuổi từ 14-25, kể cả những người chưa lập gia đình và đã lập gia đình. Trong các cuộc thảo luận này, thanh thiếu niên đã đóng góp ý kiến về thuật ngữ quen thuộc mà các em thường dùng, cách thức đặt câu hỏi, trình tự và tính liên tục của các câu hỏi cũng như các câu hỏi mà họ muốn tự điền hơn là được các điều tra viên hỏi trực tiếp. Kết quả là một số câu hỏi được sửa lại để có ngôn ngữ và cách hỏi phù hợp, một số câu được đưa vào phần tự điền.

Trước khi thử nghiệm câu hỏi ở thực địa, một lớp tập huấn đã được tổ chức cho các cán bộ từ các Cục Thống kê các tỉnh Tuyên Quang, Huế, thành phố Hồ Chí Minh đại diện cho các vùng Bắc, Trung, Nam. Khoảng 50 thanh niên nam, nữ trong và ngoài trường học, chưa lập gia đình và đã lập gia đình tham gia vào buổi thực hành phỏng vấn thử. Những thanh thiếu niên này bày tỏ cảm nghĩ của họ về cuộc phỏng vấn, các câu hỏi, những điều họ thích và không thích về quá trình phỏng vấn, việc sắp xếp chỗ ngồi, cho ý kiến về các chủ đề mà họ cho là còn thiếu, thế nào là một người phỏng vấn tốt, v.v... Công tác thử nghiệm phiếu hỏi ở thực địa được thực hiện với khoảng 180 thanh thiếu niên 14-25 tuổi từ danh sách các hộ đã tham gia điều tra mức sống tại 6 xã, phường của 3 tỉnh/thành phố nói trên.

**Giai đoạn 2:** bao gồm việc phát triển thêm các phần trong phiếu hỏi do Tổng cục Thống kê điều phối. Tại cuộc họp rút kinh nghiệm việc thử nghiệm phiếu hỏi ở thực địa, các chuyên gia đã đề xuất thử nghiệm một lần nữa để có thêm kinh nghiệm về việc phỏng vấn thanh thiếu niên ở đô thị và thanh thiếu niên dân tộc thiểu số qua phiên dịch. Sau khi thử nghiệm phiếu hỏi lần thứ 2 ở Hà Nội và Yên

Bái, các ý kiến phản hồi được tổng hợp để sửa chữa lại phiếu hỏi và sử dụng tại khóa tập huấn nghiệp vụ cho điều tra viên. Tại đây, những ý kiến đóng góp cũng được tiếp tục bổ sung vào bộ phiếu hỏi và hoàn thiện trước khi triển khai điều tra thực tại các địa bàn.

Phiếu hỏi gồm hai phần: Phần phỏng vấn và phần tự điền với tổng số khoảng hơn 200 câu hỏi. Chuyên gia SAVY đã chỉnh sửa phiếu hỏi sao cho các câu hỏi thật dễ hiểu và tránh sử dụng các thuật ngữ chuyên môn. Phần phỏng vấn bao gồm các câu hỏi chung được phân loại theo các chủ đề khác nhau, phần tự điền gồm 52 câu hỏi nhạy cảm mà thanh thiếu niên muốn trả lời một cách riêng tư. Lúc đầu phần tự điền dự kiến khoảng 10-15 câu hỏi, nhưng sau đó được phát triển thêm vì có nhiều câu hỏi ở phần phỏng vấn thanh thiếu niên cho rằng rất nhạy cảm và nên đưa vào phần tự điền. Hoàn thành xong bộ phiếu hỏi hết khoảng 60-80 phút (có thể dài hơn với đối tượng không biết đọc tiếng Việt).

Nội dung phiếu hỏi được chia thành các nhóm chủ đề sau:

- Thông tin về cá nhân và gia đình
- Học tập và trình độ học vấn
- Học nghề và việc làm
- Dạy thì/kiến thức và hành vi sức khỏe sinh sản
- Hẹn hò và tình bạn
- Quan hệ tình dục
- HIV/AIDS
- Tai nạn thương tích; bệnh tật và sức khỏe thể chất
- Thái độ, thói quen và hành vi
- Các yếu tố xã hội và cảm xúc
- Thông tin đại chúng
- Mong ước tương lai

(Xem thêm phiếu hỏi ở phần phụ lục)

### 2.3. Phương pháp điều tra

Thông tin được thu thập bằng cách mời thanh thiếu niên đến một địa điểm công cộng tại địa phương (trụ sở ủy ban, nhà văn hóa xã thôn...). Hình thức bố trí bàn ghế tại địa điểm phỏng vấn đảm bảo tính riêng tư khi phỏng vấn. Tuy nhiên, một số khó khăn trong việc vận động thanh thiếu niên tham gia phỏng vấn là không thể tránh khỏi, làm giảm số lượng thanh thiếu niên tham gia điều tra. Để khắc phục điều này, Tổng cục Thống kê đã phối hợp chặt chẽ với Đoàn

Thanh niên Cộng sản Hồ Chí Minh, Hội Liên Hiệp Phụ Nữ Việt Nam và chính quyền địa phương thông báo mục đích ý nghĩa cho các đối tượng thuộc diện điều tra, tích cực huy động tối đa thanh thiếu niên trong địa bàn được chọn tham gia.

Thanh thiếu niên và người phỏng vấn cùng giới ngồi cạnh nhau trong quá trình phỏng vấn (phương pháp này đã được đánh giá tốt qua điều tra thử). Phỏng vấn trực tiếp sẽ giúp những người được phỏng vấn làm quen với phiếu hỏi. Qua phỏng vấn trực tiếp, điều tra viên sẽ biết trình độ văn hóa của đối tượng được phỏng vấn và khả năng làm phần tự điền. Sau khi xong phần phỏng vấn, điều tra viên hướng dẫn cụ thể cách làm phần tự điền trước khi trao phiếu hỏi cho người được phỏng vấn.



Sau khi hoàn thành phiếu hỏi, người trả lời tự mang phiếu hỏi bỏ vào hòm phiếu. Đây là một thủ tục nhằm đảm bảo tính riêng tư của thông tin được thu thập.

## 2.4. Tập huấn nghiệp vụ và thu thập thông tin tại địa bàn

Ban Điều hành cuộc điều tra được thành lập với các thành viên của Tổng cục Thống kê, Bộ Y tế, Tổ chức Y tế Thế giới, UNICEF, Đoàn Thanh niên Cộng sản Hồ Chí Minh và Hội Liên hiệp Phụ nữ Việt Nam.

Vì đây là cuộc điều tra xã hội học tìm hiểu thông tin về nhiều vấn đề khác nhau, kể cả những vấn đề nhạy cảm, việc tập huấn nghiệp vụ cho cán bộ điều tra thực địa là rất quan trọng nhằm đảm bảo chất lượng cuộc điều tra. Tổng cục Thống kê đã tổ chức 3 lớp tập huấn, mỗi lớp 5 ngày, bao gồm 1 ngày thực hành phỏng vấn thử cho 150 điều tra viên từ Tổng cục Thống kê và các Cục Thống kê ở các tỉnh, thành phố.

Nội dung tập huấn: Kỹ năng, phương pháp giao tiếp với thanh thiếu niên, nội dung chi tiết của phiếu hỏi, vấn đề cần nhấn mạnh cho từng câu hỏi, cách ghi thông tin cho từng câu hỏi. Phương pháp tập huấn bao gồm hướng dẫn của giảng viên, thảo luận nhóm với thực hành phỏng vấn thử. Có hơn 150 thanh thiếu niên thuộc các nhóm xã hội khác nhau đã được mời đến để các học viên tiến hành phỏng vấn thử. Kết thúc mỗi cuộc phỏng vấn thử, các thanh thiếu niên này được đề nghị cho biết ý kiến/cảm nghĩ của mình về nội dung được hỏi và cách phỏng vấn của học viên. Những đóng góp của họ rất có ý nghĩa đối với các giảng viên và bản thân học viên về phương pháp phỏng vấn, kỹ năng làm việc với thanh thiếu niên. Chỉ có những học viên hoàn thành tốt bài kiểm tra sau lớp học mới được tham gia cuộc điều tra.

Điều tra thực địa diễn ra trong 53 ngày. Điều tra viên được chia làm 19 đội. Trước cuộc điều tra, các đội chuẩn bị lịch trình đi lại cụ thể giữa các địa bàn, các huyện và vùng. Thông qua Hội Liên hiệp Phụ nữ, Đoàn Thanh niên, Ủy ban nhân dân xã, phường, thị trấn thông báo cho thanh thiếu niên và huy động họ tham gia vào cuộc điều tra.

## 2.5. Sai số trong điều tra và hạn chế của số liệu

Các số liệu trong báo cáo này là những chỉ số ước lượng lấy từ SAVY. Có 2 dạng sai số: sai số chọn mẫu và sai số phi chọn mẫu. Sai số chọn mẫu xảy ra do các quan sát được thực hiện trên mẫu chứ không phải tổng thể. Sai số phi chọn mẫu có thể do nhiều lý do khác nhau. Độ chính xác của điều tra chọn mẫu sẽ được xác định bởi tác động phối hợp của hai loại sai số này. Với điều tra sử dụng kỹ thuật chọn mẫu xác suất, có thể đánh giá sai số chọn mẫu của bất kỳ kết quả ước lượng cụ thể nào bằng cách sử dụng chính số liệu trong nghiên cứu khảo sát.

## 2.6. Sai số phi chọn mẫu

Sai số phi chọn mẫu không dễ lượng hóa như sai số chọn mẫu vì nó rất đa dạng, mỗi kiểu sai số phi chọn mẫu đều phải được đánh giá riêng để xác định ảnh hưởng của nó tới kết quả điều tra. Có thể điểm qua một vài sai số phi chọn mẫu: (1) hộ gia đình hoặc thanh thiếu niên được chọn nhưng không tham gia điều tra, (2) khó khăn trong khái niệm, định nghĩa trong quá trình thiết kế, (3) về phía người được hỏi không thể hoặc không muốn cung cấp thông tin chính xác, (4) điều tra viên và người nhập số liệu mắc lỗi trong việc đánh mã, nhập mã, (5) các lỗi khác khi thu thập thông tin, xử lý và báo cáo. Việc đánh giá sai số phi chọn mẫu còn hạn chế ở chỗ lý thuyết thống kê về vấn đề này chưa được phát triển như ở sai số chọn mẫu và ngoài ra còn thiếu kiến thức về các giá trị thực trong dân số nghiên cứu. Vì những lý do này, những người thiết kế và thực hiện những cuộc điều tra với quy mô lớn như SAVY, hiếm khi cung cấp những thông số liên quan đến thể loại và tầm cỡ của các sai số phi chọn mẫu có thể có. Ngoài ra, ngân sách hạn chế cũng làm hạn chế tính khả thi.

Việc có thể làm để khống chế và giảm thiểu sai số phi chọn mẫu là biên soạn bảng hỏi kỹ lưỡng, điều tra thử để kiểm tra bảng hỏi, phương pháp và công tác tổ chức, tập huấn bài bản, kiểm tra điểm tại thực địa, kiểm định nhập liệu và mã hóa, giám sát chặt chẽ. Tất cả các bước trên đều được thực hiện trong SAVY. Tuy nhiên vẫn cần phải mô tả sai số phi chọn mẫu hoặc thiên lệch (biases) nếu có và chỉ ra tác động có thể của chúng.

Một nguồn sai số phi chọn mẫu quan trọng của SAVY là tỷ lệ không tham gia. Cách tính bằng  $(1 - I/n)$ , trong đó  $I$  là số lượng thanh thiếu niên 14-25 được phỏng vấn thực và  $n$  là con số được chọn trong mẫu. SAVY có 7.584 thanh thiếu niên được phỏng vấn trên 9.989 được chọn, vì vậy tỷ lệ không tham gia =  $(1 - 7.584/9.989) = 24,1\%$ . Ở mức độ này, việc không tham gia có thể có tác động sai lệch đáng kể tới kết quả bởi lẽ ta giả định số không tham gia cũng có những đặc điểm và sự phân bố tương tự như số tham gia điều tra. Giả định này không thể được kiểm chứng một cách độc lập.

So sánh phân bố mẫu giữa nghiên cứu này với phân bố mẫu có từ các nguồn khác, đặc biệt là từ các nghiên cứu khác và điều tra y tế quốc gia cho

thấy sai lệch là không đáng kể, trừ một ngoại lệ: Mẫu của SAVY hơi thiếu hụt nhóm thanh niên thuộc nhóm tuổi lớn, đó là những thanh niên đi lấy chồng, đi học, đi làm xa nhà. Điều này thể hiện rõ khi so sánh nhóm độc thân, đi học, đi làm chia theo tuổi và giới tính của điều tra SAVY với tổng điều tra dân số 1989 - 1999. Sai lệch này không ảnh hưởng lớn tới việc phân tích số liệu SAVY, nhưng dù sao cũng cần lưu ý.

## 2.7. Các sai số chuẩn

Mẫu được chọn trong SAVY là một trong các mẫu có cùng cỡ có thể được chọn ngẫu nhiên. Giá trị đặc thù của mỗi ước lượng được gọi là ước lượng điểm và các ước lượng điểm lấy từ mỗi mẫu lại khác nhau. Sự chênh lệch của ước lượng mẫu so với giá trị trung bình của các ước lượng điểm được gọi là sai số chọn mẫu. Không thể tính toán chính xác được sai số chọn mẫu thực sự vì ta chỉ làm trên một mẫu, sai số chuẩn của một ước lượng như được tính toán trong báo cáo này chính là ước lượng của sai số chọn mẫu. Sai số chuẩn ước tính đồng thời cũng đo được một phần ảnh hưởng của sai số phi chọn mẫu, ví dụ như sai số do sự khác nhau giữa các phỏng vấn viên, người nhập mã nhưng không thể đo được những sai lệch hệ thống trong số liệu.

Ước lượng điểm trong mẫu cho một biến số hoặc chỉ số nào đó và ước lượng điểm sai số chuẩn cho phép ta xây dựng được khoảng tin cậy chứa các ước lượng điểm của các mẫu (nếu được chọn). Để minh họa, nếu ta chọn tất cả các mẫu, tiến hành điều tra trong cùng một điều kiện, tính sai số chuẩn cho từng mẫu, thì xấp xỉ 95% của các khoảng từ 2 lần sai số chuẩn dưới đến 2 lần sai số chuẩn trên sẽ chứa tất cả các giá trị trung bình của các mẫu và được gọi là khoảng tin cậy 95%.

Chi tiết về một số sai số sẽ được trình bày trong phần phụ lục.

1. Vào thời điểm tiến hành điều tra, trên cả nước có 61 tỉnh/thành phố tuy nhiên vào thời điểm công bố báo cáo, cả nước có 64 tỉnh/thành phố.