

Sławomir Pasikowski

Akademia Pomorska w Słupsku, e-mail: s.pasikowski@apsl.edu.pl

Czy wielkość jest niezbędna? O rozmiarze próby w badaniach jakościowych

<http://dx.doi.org/10.12775/PBE.2015.055>

Is The Size is Necessary? About The Sample Size in Qualitative Research

Abstract

The question about a sample size in qualitative research has been marginalized. Essentially because the issue is treated as peripheral and dependent on the state of theoretical saturation. The article provides an overview of research on the sample size coinciding with achieving this state. The conditions of the saturation and the factors determining the sample size also are discussed. On the base of the solutions developed in statistical research area the conception of mixing methods of sampling to maximize the credibility of the data and the theoretical saturation has been presented.

Key words: representativeness; sample size; saturation; theoretical sampling; transferability

Panuje przekonanie, że badania jakościowe tym różnią się od ilościowych, że te ostatnie skupiają się na dużych próbach, podczas gdy w jakościowych badania prowadzi się na nielicznej liczbie przypadków, czasem nawet jednym. Nie do końca jednak przekonanie to jest zgodne z istniejącym stanem rzeczy. W badaniach ilościowych dopuszcza się analizy prowadzone na małych próbach, co więcej z możliwością jednoczesnego uzyskiwania reprezentatywności przy znajomości przynajmniej takich warunków, jak poziom ufności, zmienność cechy oraz zakładany błąd szacunku. Z kolei w jakościowych liczba zdarzeń lub

podmiotów poddawanych obserwacji i analizie osiąga rozmiary wielokrotnie przekraczające wartość kilkunastu.

To, że niewiele zdarzeń lub przypadków potrzeba, aby móc w sposób akceptowalny realizować badanie jakościowe, jest stosunkowo oczywiste w środowisku ludzi skupionych wokół nauk społecznych. Wątpliwości jednak występują wówczas, gdy próba stale się rozszerza i brakuje kryterium określającego moment ostatecznego zakończenia zbierania materiału empirycznego. Sytuacja taka może zatem prowokować pytanie o rozmiar próby w badaniach jakościowych i o moment, do którego możliwe jest sensowne poszerzenie jej zakresu o kolejne przypadki. Spotkać się można z raportami z badań, w których liczba jednostek¹ lub indywidualów poddawanych obserwacji wyrażana jest w setkach. Jeśli metodą gromadzenia materiału był wywiad narracyjny, to nietrudno wyobrazić sobie ogrom pracy, przed jakim stawał badacz, nawet wtedy, gdy samej analizie poddawane były tylko nieliczne elementy z każdego z wywiadów. Taki wysiłek skłania niewątpliwie do pytania nie tylko o zasadność takiej liczby wywiadów, ale też na ile analizy i wnioski z nich wyprowadzane zasługują na zaufanie.

Każde badanie naukowe obciążone jest ryzykiem błędnego rezultatu. To stwierdzenie należy do kanonu podstawowych zasad metodologicznych (Kamiński, 1989; Taylor, 1995; Hajduk, 2011). Na ogólny błąd badania składają się błędy generowane przez badacza, badane jednostki, metodę pomiaru i analizy danych oraz ogólny sposób organizacji procesu badania. Świadomość tych ograniczeń jest przyczyną dyskusji na temat rygoru metodologicznego w badaniach jakościowych, która swym zasięgiem obejmuje także rozmiar próby i sposób jej pozyskiwania (por. Guba, 1981; Sandelowski, 1993; Morse i in., 2002; Marshall i in., 2013). Wielkość próby może zarówno stanowić źródło błędów, jak i zabezpieczać przed błędnymi rezultatami badawczymi. Niekiedy więc mały jej rozmiar może ograniczać uchwycenie charakteru zjawiska. Jednak nawet bardzo liczna próba nie zastąpi starannie zaplanowanego doboru, który podnosi jakość uzyskiwanych danych, również wówczas, gdy rozmiar próby jest zdecydowanie niewielki. W związku z tym podczas dyskusji na temat rozmiaru próby w badaniach jakościowych na pierwszy plan wysuwa się teza o priorytetowym znaczeniu warunków, w jakich zachodzi jej dobór.

¹ Pojęcie „jednostka”, w zależności od kontekstu używane jest w tekście w znaczeniu „osoba” lub „element zbioru”, np. zbioru zachowań, motywacji, reguł, rytuałów, emocji, grupy społecznej, dowolnego rodzaju *social units*. W tekście synonimicznie stosowane jest też pojęcie „przypadek”.

Rozmiar próby w teorii i praktyce badań jakościowych

Rozmiar próby w badaniach jakościowych, mimo że traktowany jako mający drugorzędne znaczenie, jest jednak przedmiotem pytań o charakterze zarówno praktycznym, jak i teoretycznym. Przyjmuje się najczęściej, że o wielkości próby w badaniach jakościowych decyduje stan teoretycznego nasycenia generowanych kategorii, a który polega na replikowalności informacji uzyskiwanych w trakcie pracowania z materiałem empirycznym (Konecki, 2000; Ritchie, Lewis, Elam, 2003; Cohen, Manion, Morison, 2007; Creswell, 2007, s. 240; Charmaz, 2009; Glaser, Strauss, 2009; Thomson, 2011). Generalnie też samo zagadnienie wielkości próby traktowane jest marginalnie, co nie powinno dziwić, gdy podstawowym dążeniem w badaniach jakościowych czyni się budowanie teorii możliwie bezpośrednio z dostępnych danych. Nie zmienia to jednak tego, że badacze jakościowi zgłaszają potrzebę określania wielkości próby (Baker, Edwards, 2012). Samo pojęcie nasycenia teoretycznego nie zaspokaja oczekiwań, gdyż w zasadzie sprowadza się do abstrakcyjnej idei punktu, w którym nowe dane nie wnoszą już nic albo wnoszą niewiele znaczących zmian do bieżącej teoretycznej rekonstrukcji zjawiska (por. Charmaz, 2009, s. 141; Guest, Bunce, Johnson 2009, s. 65). Dzieje się tak przede wszystkim dlatego, że osiągnięcie tego stanu okazuje się wypadkową wielu czynników (Guest i in., 2009; Marshall i in., 2013). Zależy bowiem nie tylko od ilości i kompletności danych, ale też od indywidualnej charakterystyki badacza oraz liczby osób uczestniczących w procesie analizy. Wiedza analityka, jego doświadczenie, styl poznawczy oraz bieżąca kondycja intelektualna determinują ilość informacji dostrzeganych w surowych danych, liczbę generowanych kodów i kategorii. Te indywidualne właściwości same tworzą więc kontekst dla generowania treści z dostępnego materiału empirycznego. Wskazane uwarunkowania decydują o przypisywaniu teoretycznemu nasyceniu etykiety stanu opartego w dużym stopniu na subiektywizmie, niejasnych przesłankach i bliżej nieokreślonej treści (Marshall i in., 2013; O'Reilly, Parker, 2013). W związku z tym i rozmiar próby może wydawać się powzięty zupełnie arbitralnie, gdy badacz wyłącznie wzmiankuje o zastosowaniu doboru opartego na kryterium nasycenia i nie podaje, w jaki sposób oraz kiedy stan ten został osiągnięty. Wyniki prowadzonych obserwacji pokazują, że niestety często tak właśnie bywa (Bowen, 2008; Marshall i in., 2013).

Problem wielkości próby trudno też uznać za rozstrzygalny przy pomocy idei teoretycznego nasycenia. Przede wszystkim z powodu proceduralnej kontrowersji, jaka wiąże się z tą ideą. Pojęcie teoretycznego nasycenia nie pozwala bowiem przybliżyć, w jakim stopniu doświadczenia lub właściwości związane

z badanym zjawiskiem zostały objęte doborem. Poza tym o momencie osiągnięcia nasycenia decyduje sam badacz, i to nierzadko według niejasnych kryteriów (Bowen, 2008). Zaproponowano w związku z tym łagodniejszy termin: teoretyczna wystarczalność (Day, 1999, za: Charmaz, 2009). Jednakże dał on odpór jedynie zarzutom nieuzasadnionej pewności co do osiągnięcia nasycenia. Pozostawił więc bez odpowiedzi pytanie o choćby przybliżoną granicę zaprzestania doboru kolejnych jednostek do próby.

Wskazówek, dających wyobrażenie w tej kwestii, udzielają badania nad wielkością próby w projektach jakościowych. Ujawniają one, że w praktyce przeciętna wielkość, na jaką decydują się badacze, mieści się w granicach 20–30 jednostek. Jednakże w określeniu satysfakcjonującego dla badacza rozmiaru próby decydujące znaczenie ma zastosowana metoda oraz schemat badania. Prowadzone analizy nad wielkością próby w projektach jakościowych (Creswell 2007, s. 12; Mason, 2010) wykazują, że w badaniach etnograficznych realizowanych jest zwykle 30–60 wywiadów w ramach jednego projektu, w badaniach wspartych na założeniach fenomenologii lub fenomenografii wykorzystuje się najczęściej 5–25 wywiadów, natomiast realizacja badań w nurcie teorii ugruntowanej najczęściej zbiega się ze stosowaniem próby o rozmiarze 20–30 jednostek. Istnieją też dobrze udokumentowane dane mówiące o osiągnięciu nasycenia przy 6–12 przypadkach, jak też stabilizowaniu się liczby i treści kodów oraz zbieżności materiału pozyskiwanego z wywiadów już przy 24 przypadkach (Guest i in., 2006). W prowadzonych przez Mark Mason (2010) rozległych badaniach obejmujących blisko 600 odrębnych jakościowych projektów empirycznych rozmiar próby wahał się najczęściej w granicach 20–50 przypadków, przy czym nigdy nie przekraczał liczby 95. Materiał zorganizowany w badaniach nad polskimi projektami (Pasikowski, 2014a; 2014b) dał w efekcie podobne rezultaty do tych, jakie uzyskał Mason. Jednak w polskiej próbie raportów około 10% (blisko 8 raportów) stanowiły te, które opierały się na próbach liczących ponad 100 jednostek. W jednym z nich informowano o próbie liczącej 374 przypadki. Wprawdzie przy projektach jakościowych taka liczebność nie wydaje się zasadniczym błędem, to jednak w kontekście idei nasycenia teoretycznego oraz wysiłku, jaki należy włożyć w gromadzenie i analizę materiału, świadczy raczej o nieliczeniu się z czasem oraz zasobami ekonomicznymi (por. Adler, Adler, 2012, s. 10).

Spotkać się można z argumentacją, że odpowiednie przygotowanie doboru do próby może dawać w efekcie wiarygodne dane już przy niezwykle małych próbach, jak wtedy gdy liczba jednostek wynosi 4 (Romney, Batchelder, Weller, 1986, za: Guest i in., 2009, s. 74). Badani o wysokim poziomie kompe-

tencji kulturowych odnoszących się do kontekstu kulturowego, którego sami są uczestnikami, dostarczać mieliby daleko bardziej wiarygodnych informacji niż liczna próba osób pozostających na marginesie tego kontekstu. Takie stanowisko odsłania jednak konieczność brania pod uwagę dwóch zasadniczych warunków decydujących o wielkości próby. Pierwszym jest cel badania, a drugim znajomość populacji i jej kontekstu. Jeśli badacza interesują indywidualne cechy przedstawicieli populacji, to wymagana wielkość próby do osiągnięcia wystarczającego poziomu informacyjnej zawartości danych będzie raczej większa niż mniejsza. Podobnie w przypadku drugiego warunku, wymagana wielkość próby będzie tym większa, im mniej informacji będzie posiadał badacz o przedstawicielach danej populacji.

Greg Guest, Arwen Bunce i Laura Johnson (2009, s. 79) przekonują, że jeśli celem badacza jest zrozumienie spostrzeżeń i doświadczeń podzielanych przez przedstawicieli jakiejś względnie homogenicznej populacji, a sama dziedzina, która jest eksplorowana, jest jednocześnie względnie spójna i przejrzysta, to próba składająca się z kilkunastu jednostek powinna okazać się satysfakcjonująca. Sugestia ta nie wydaje się jednak przekonująca. Propozycja nie jest przez autorów dostatecznie uzasadniona, a dostarczona przez nich argumentacja nie pozwala też sformułować odpowiedzi na następujące pytania: na jakim podstawie badacz może zakładać, że populacja, którą bada jest homogeniczna? jak określić kryteria tej homogeniczności? czy przypadki zakwalifikowane do próby wystarczająco dobrze reprezentują właściwości populacji lub samo badane zjawisko? jak ocenić właściwości zjawiska zanim podda się je obserwacji? oraz jeśli obraz populacji budowany będzie w toku gromadzenia kolejnych obserwacji, to co będzie uwiarygodniało traktowanie wyłaniającej się wizji jako ilustracji charakterystyki szerszej zbiorowości, a przecież opartej na informacjach dostarczanych przez wybrane jednostki?

Można wprawdzie jeszcze bronić stanowiska Guest i współautorów, powołując się na ideę podzielanych przekonań kulturowych. Zwykle bowiem idea ta uchodzi za wystarczające uzasadnienie wiarygodności danych z prób wygodnych i małowyczących. Zgodnie z tą ideą przyjmuje się, że osoby biorące udział w badaniu zatopione są w szerszym kontekście kulturowym przez to, że przynależą do grup, zbiorowości lub społeczeństw, które same pozostają w zasięgu tego kontekstu, podtrzymują go i współtworzą. Stosowanie tego argumentu nie uwzględnia jednak, że w momencie gromadzenia materiału empirycznego badacz nie zawsze wie, jakie są przekonania kulturowe danej populacji oraz czy dobrana grupa uczestników podziela te przekonania w podobnym stopniu i zakresie, co reszta zbiorowości. Zbieżność rezultatów 6, 12 lub nawet 30 pierw-

szych przypadków może być wynikiem nieuświadomianego przez badacza podobieństwa w zakresie innych cech niż te, które jako świadome kryterium zostały przyjęte podczas doboru próby i które wcale nie muszą być typowe dla populacji będącej adresatem dokonywanych później uogólnień z badań. Powoływanie się podczas doboru do próby na prymarne znaczenie podzielanych przekonań kulturowych może więc mieć więcej wspólnego z metodologiczną naiwnością niż z metodologiczną wiarygodnością.

Podsumowując, żadna ze wskazanych wielkości próby w przywołanych badaniach nie została merytorycznie uzasadniona. Są one albo efektem arbitralnej decyzji indywidualnych projektantów badań podpartej powoływaniem się na konstrukt nasycenia teoretycznego albo sugerowaniem się wartościami, na jakie decyduje się większość badaczy. Taka sytuacja czyni bardziej zrozumiałym poszukiwanie przez niektórych badaczy merytorycznych wskazówek określających granice doboru próby w badaniach jakościowych.

Reprezentatywność i idea niezbędnej wielkości próby

W badaniach reprezentacyjnych, czyli opartych na filozofii rachunku prawdopodobieństwa i idei reprezentatywności próby względem populacji, kwestia rozmiaru próby precyzowana jest przez pojęcie „niezbędnej wielkości próby”. Termin ten odnosi się do rozmiaru próby, który jest zoptymalizowany ze względu na możliwość formułowania uogólnień z pobranej próby na populację, z której ta próba pochodzi. Wielkość takiej próby jest zminimalizowana, jednak w stopniu umożliwiającym generalizowanie rezultatów badawczych. Warunkami określenia tak rozumianego minimalnego rozmiaru próby jest znajomość przynajmniej zmienności cechy będącej przedmiotem zainteresowania badacza, poziomu ufności, na jakim prowadzone ma być późniejsze wnioskowanie statystyczne oraz określenie dopuszczalnego błędu szacowania charakterystyki tej cechy w populacji na podstawie charakterystyki występującej w próbie. Czynnikiem użytecznych i możliwych do uwzględnienia w określaniu rozmiaru próby jest jednak znacznie więcej (por. Ritchie, Lewis, Gillian, 2003; Szreder, 2010).

Łatwo jest zauważyć, że podstawą wyprowadzenia niezbędnej wielkości próby jest skupienie na cesze lub cechach populacji, a nie na jej jednostkach. Zasadne jest więc pytanie: czy wzorem badań reprezentacyjnych o niezbędnej wielkości próby w badaniach jakościowych może być w ogóle mowa? Odpowiedź nie jest bezwzględnie negatywna. Po pierwsze, w badaniach jakościowych dopuszcza się tworzenie uogólnień w oparciu o uzyskane rezultaty ba-

dawcze (Silverman, 2008, s. 274; Flick, 2011a, s. 13; Gibbs, 2011, s. 13, 27), co też nazywane jest transferowalnością (*transferability*) i uchodzi za odpowiednik trafności zewnętrznej w badaniach statystycznych (Guba, 1981; Lewis, Ritchie, 2003); formułowanie uogólnień następuje także wtedy, gdy budowane są klasyfikacje obiektów. Po drugie, koncepcje teoretycznego nasycenia i doboru teoretycznego odnoszą się bezpośrednio do idei rekonstrukcji zróżnicowania, ale też odkrywania tego, co typowe i regularne, a zatem w jakiś sposób reprezentatywne dla danego obszaru, kategorii osób lub zjawiska (por. Glaser, Strauss, 2009, s. 53; Flick, 2011a, s. 13; 2011b, s. 61). Po trzecie, dobór teoretyczny wraz z towarzyszącym mu teoretycznym nasyceniem stanowi jedną ze strategii weryfikacyjnych w badaniach jakościowych, które zabezpieczają przywołaną przed momentem transferowalność (Guba, 1981; Morse i in., 2002). To właśnie w teoretycznym nasyceniu upatrywana jest podstawa tworzenia uogólnień w badaniach niestatystycznych (Gorzko, 2001, s. 71). Wszystko to sprawia, że dla badacza jakościowego znacząca może być nie tylko reprezentatywność próby lub reprezentatywność danych, jakich ta próba dostarczy. Takiego badacza może w związku z tym nurtować również rozmiar próby skojarzony z momentem osiągania reprezentatywności, rozumianej także jako dokonywanie wiarygodnej klasyfikacji. Tym bardziej, że pojęcie reprezentatywności bezpośrednio powiązane jest z zagadnieniem wielkości próby. To bowiem spełniający odpowiednie kryteria merytoryczne poziom reprezentatywności odnoszony do zjawiska lub populacji stanowi granicę doboru.

Ostatnie stwierdzenie jest jednak powinnościowym postulatem metodologicznym. Praktyka badawcza dowodzi, że również badaniami statystycznymi mogą kierować bliżej nieokreślone kryteria doboru próby oraz jej rozmiaru. Bywa, i to wcale nie rzadko, iż o wielkości próby decyduje przekonanie, zgodnie z którym im więcej, tym wiarygodniej, bo tym wierniej wyniki badań reprezentują badane zjawisko. Metodologia badań reprezentacyjnych przeczy temu i pozwala ufać, że przy zastosowaniu wybranej formuły obliczeniowej oraz odpowiadającej jej selekcji jednostek już dwucyfrowa liczba przypadków może być stanowczo bardziej reprezentatywna względem rozległej populacji niż próba licząca nawet kilka setek jednostek.

W przypadku projektów jakościowych zdarzają się i takie, w których badacze zainteresowani są wyłącznie jednym konkretnym przypadkiem. Oni jednak także stosują dobór według jakiegoś kryterium. Takich badaczy znacznie mniej może interesować reprezentatywność lub transferowalność. Choć nie można wykluczyć, że staną w pewnym momencie przed dylematem oceny zachowań, zdarzeń lub właściwości, w zbiorze tych wszystkich charakteryzujących bada-

ny przypadek, które są typowe, dominujące, najczęściej występują lub występują z największą intensywnością. A zatem takich, które wyznaczają specyfikę tego właśnie przypadku.

Podręczniki metodologii badań społecznych (np. Babbie, 2006; Cohen i in., 2007; Nowak, 2007; Lohr, 2010) uczą, że próba jest reprezentatywna, jeśli spełnia warunek takich samych szans trafienia jednostek z populacji do próby lub szans odpowiadających poziomowi występowania jednostek określonego typu w tejże populacji. Dzięki temu próba będzie nieistotnie różniła się od populacji, w zakresie cech będących przedmiotem badania. Przy takim ujęciu trudno bezwzględnie zgodzić się z podręcznikową wytyczną. Wiadomo przecież, że aspiracje do organizowania próby oddającej cechy szerszej zbiorowości charakteryzują nie tylko statystycznie zorientowanych badaczy. Przykładem na to jest spójne stanowisko metodologiczne w postaci teorii ugruntowanej, które za cel stawia sobie formułowanie uogólnień na podstawie uzyskiwanych rezultatów badawczych (Gorzko, 2001; Glaser, Strauss, 2009). Wspomnieć przy tym warto o atrakcyjnym dla badaczy jakościowych pojęciu reprezentatywności typologicznej, która polega na odzwierciedlaniu wszystkich wariantów zjawiska i ich kombinatoryki (Gorzko, 2001; Nowak, 2007, s. 300).

W zabieganiu o reprezentatywność nie chodzi jednak o zorganizowanie próby, która stanowi miniaturę populacji, czyli odzwierciedla wszystkie cechy populacji wraz z ich charakterystyką i strukturą. Tak pojęta reprezentatywność nawet w badaniach statystycznych uchodzi za niewiarygodną, a idea próby będącej miniaturą kopią populacji identyfikowana jest z tzw. mitem próby reprezentatywnej (Siekierski, 2002). Zwykle zabiega się o reprezentatywność w zakresie wybranych cech, które są ważne dla celów badania oraz stawianych pytań. Co więcej, traktuje się reprezentatywność jako przybliżoną zbieżność rozkładu cechy w próbie z rozkładem tej samej cechy w populacji. Oznacza to, że zakres wartości lub wariantów cechy nie musi być identyczny. Choć taki właśnie byłby preferowany.

Reprezentatywność bywa też definiowana jako zróżnicowanie znaczenia, „jake badane zjawisko ma w świadomości i praktyce życiowej badanych osób” (Flick, 2011b, s. 61). Mowa jest wówczas o odzwierciedlaniu znaczeń w świadomości osób, a nie o odzwierciedlaniu cech populacji. Gdy jednak przyjrzeć się bliżej temu konstruktowi, to okazuje się, że nie stoi on w opozycji do klasycznie pojmowanej reprezentatywności. Możemy przecież mówić np. o populacji znaczeń, jakimi dysponuje badana osoba, grupa czy nawet cała zbiorowość ludzi. Stąd łatwiej przyjąć, że to, co ulega zmianie (czyli desygnat populacji), nie implikuje konieczności modyfikowania technik doboru materiału empirycznego.

To badacz definiuje populację oraz dobiera instrumenty gromadzenia i analizy danych stosownie do celu badania, jak też odpowiednio do pytań, jakie stawia.

Wyznaczanie czynników określających niezbędną wielkość próby w badaniach jakościowych jest niewątpliwie dyskusyjne. Wprawdzie podczas ustalania rozmiaru próby można kierować się wynikami przytaczanych badań. Ale w tego rodzaju badaniach, co trzeba mocno podkreślić, przecież nie o wielkość próby chodzi, lecz o jej możliwości w zakresie dostarczania wiarygodnych informacji na temat zjawisk i właściwości charakteryzujących zbiorowość, z której ta próba pochodzi. Liczy się więc dążenie do podnoszenia stopnia wysycenia danych empirycznych informacjami o właściwościach zbiorowości, którą badacz się interesuje. Jest to możliwe poprzez intensyfikowanie nasycenia teoretycznego kategorii i maksymalizowanie przynajmniej typologicznej reprezentatywności próby.

Podnoszenie zawartości informacyjnej danych: łączenie doboru teoretycznego z metodą reprezentacyjną

Dobór do próby kierowany ideą nasycenia teoretycznego tworzonych kategorii nazywany jest doбором teoretycznym. Nie służy jednak zasadniczo gromadzeniu surowego materiału, choć ten jest jego podstawą, co raczej dopasowywaniu wyłaniającej się teoretycznej wizji zjawiska do danych empirycznych (Konecki, 2000; Creswell, 2007; Charmaz, 2009; Glaser, Strauss, 2009). Jaśniejszy staje się charakter tego doboru, gdy zauważyć, że obejmuje on także wstępny czy też początkowy dobór próby (por. Glaser, Strauss, 2009, s. 41–42). Ten początkowy dobór dostarcza materiału, z którego wyprowadzane są wstępne kategorie teoretyczne. Kolejne doборы jednostek mają już charakter teoretyczny, gdyż kierowane są wskazaniem teorii wyłaniającej się z dotychczas zgromadzonych danych. W każdym kolejnym doborze do już istniejących kryteriów dokładane są kolejne ograniczające swobodne dostawanie się jednostek zbiorowości do tworzonej próby (por. Konecki, 2000, s. 31; Ritchie i in., 2003, s. 81; Glaser, Strauss, 2009).

Pieczołowicie zorganizowanej próbie powstałej w efekcie zastosowania doboru teoretycznego i kryterium teoretycznego nasycenia blisko jest do tzw. próby hazardowej (por. Steczkowski, 1995; Hall, Herron, Pierce, 2006). O próbie hazardowej mówi się wówczas, gdy jednostki dobierane są według przypadku, na chybił trafił, czemu towarzyszy przekonanie o beztendencyjnym wyborze. Beztendencyjność selekcji w założeniu zapewniać ma wszystkim jednostkom zbiorowości zbliżone szanse trafienia do próby i w rezultacie

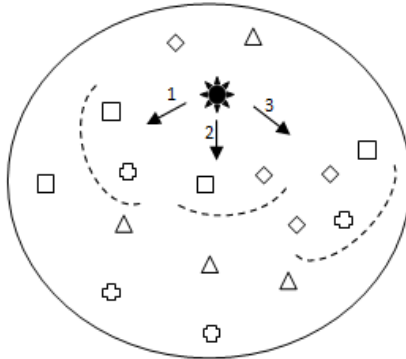
prowadzić do odzwierciedlenia cech zbiorowości w próbie. Próba teoretyczna, tak jak hazardowa, sprawia wrażenie wysoce reprezentatywnej, bo regulowanej mechanizmami względnie niezależnymi od badacza. Jednak wybieranie na chybił trafił lub kierowanie się wskazaniem bieżącej reprezentacji poznawczej zjawiska jest wysoce zdeterminowane właściwościami związanymi z charakterystyką samego badacza oraz jego usytuowania w czasie i przestrzeni (por. Steczkowski, 1995; Hall i in., 2006). O podobnym zdeterminowaniu stanu nasycenia teoretycznego mowa była już wcześniej.

Jednym z efektów tego zdeterminowania jest tzw. błąd pokrycia (Szreder, 2010). Obciążony jest nim dobór hazardowy i – o ile wcześniej nie zostały odpowiednio zabezpieczone – także dobór teoretyczny wraz z służącym mu doborem wstępnym, zwykle celowym². Błąd pokrycia należy do grupy błędów nielosowych badania i polega na ograniczeniu w dostaniu się do próby elementów reprezentujących dane grupy lub typy rzeczywiście występujące w zbiorowości. Reprezentacja właściwości populacji nie pokrywa się wówczas z realnymi właściwościami populacji. Źródło tego błędu tkwi w niekompletności wiedzy o strukturze zbiorowości, brakach w operacie stanowiącym podstawę doboru oraz w błędach technicznych doboru.

W przypadku doboru teoretycznego można ten błąd zredukować, stosując probabilistyczne techniki doboru przynajmniej w pierwszym etapie organizacji próby, czyli we wstępnym doborze³. W rezultacie dotarcie do przypadków reprezentujących przeróżne, w tym jeszcze nierozpoznane i nieświadomione, aspekty badanego zjawiska staje się bardziej możliwe. Tym samym prawdopodobniejsze jest także oddanie właściwości tego zjawiska wraz z informacyjnym wysyceniem danych i nasyceniem kategorii. Przede wszystkim badacz przestaje polegać wyłącznie na swojej intuicji, która nie zawsze gwarantuje trafny dobór próby. Rzadko też przecież wie, jak badane zjawisko lub właściwość reprezentowana jest w danej kategorii osób lub jednostek obserwacji, których badanie ma dotyczyć. Co więcej, jeśli próba nie obejmuje wszystkich przypadków danej kategorii, to tym bardziej nie może mieć pewności, że zjawisko lub cecha zostaną zrekonstruowane w kolejnych momentach postępowania badawczego.

² O wadze pokrycia w badaniach jakościowych wyraźną wzmiankę znaleźć można także w samej pracy Barneya Glasera i Anselma Straussa (2009, s. 59).

³ Łączenie doboru losowego z teoretycznym dopuszczali już Glaser i Strauss (2009, s. 54–55), choć nie rozwinęli tej idei, skupieni raczej na różnicach niż wykazywaniu możliwości wartościowej koordynacji.



Rys. 1. Ilustracja pomijania informacji w doborze kierowanym ideą nasycenia teoretycznego

Źródło: opracowanie własne.

Rysunek 1 obrazuje, w jaki sposób dochodzić może do pomijania elementów istotnych dla rekonstrukcji zjawiska. Badacz (oznaczony gwiazdką) zlokalizowany jest w obszarze danej dziedziny społecznej. Kierując się ideą doboru teoretycznego, rozpoczyna proces pobierania przypadków z miejsca, które podpowiada mu intuicja lub wiedza, jaką dysponuje. Podczas kolejnych pobrań (numerycznie oznaczone strzałki) obserwacją zostają objęte jednostki charakteryzujące się coraz to nowszymi właściwościami (symbolizują to łuki oznaczone przerywana linia). Po trzecim pobraniu badacz podejmuje decyzję o osiągnięciu stanu nasycenia kategorii z powodu braku pojawiania się nowych właściwości w ich obrębie. Oczywiście mówienie o trzech pobraniach jest pewnym uproszczeniem mającym mocniej uwidocznić dyskutowany mechanizm, gdyż w praktyce przy braku występowania nowych informacji jeszcze przez pewien czas gromadzi się dane w ramach weryfikacji uzyskanych informacji (Marshall i in., 2013). Badacz zebrał i potwierdził informacje o występowaniu trzech nieznanych dotychczas atrybutów. Jednakże dane, którymi dysponuje, nie wyczerpują możliwości, jakimi cechuje się badana zbiorowość. Istnieje bowiem czwarty atrybut. Nawet znajdowanie się w bezpośrednim sąsiedztwie obiektu cechującego się tym atrybutem nie gwarantuje objęcia go obserwacją. Z powodu niedostatków wiedzy o zbiorowości, geograficznej odległości, jaka dzieliła badacza od pozostałych przypadków lub zaistnienia innych dowolnych czynników ograniczających dostępność do danych, badacz nie osiągnął stanu, w którym możliwe byłoby pełniejsze zrekonstruowanie atrybutów zjawiska lub właściwości występujących w tej badanej zbiorowości.

W praktyce badacze próbują bronić się przed takimi sytuacjami, wydłużając proces obserwacji lub dokonując dodatkowych przybliżeń w oparciu o kolejne jednostki oraz alternatywne źródła danych i metody ich gromadzenia (Glaser, Strauss, 2009; Thomson, 2011; Marshall i in., 2013). Odnoszę jednak wrażenie, że wydłużaniu obserwacji towarzyszy niekiedy zbyt optymistyczna wiara w korygującą funkcję czasu. Miejsce, z którego się patrzy, nie musi bowiem dostarczać nowych lub teoretycznie znaczących informacji pomimo upływu czasu. Jest w tym przypadku trochę tak jak z jazdą pociągiem. Nawet jeśli przejedziemy całą trasę, to zobaczymy tylko to, co znajduje się w sąsiedztwie torów.

Z kolei korzystanie z alternatywnych źródeł i metod trudno jest przecenić. Jednakże ich możliwości można by efektywniej wykorzystywać. Szczególnie, gdy za przyczynę zredukowanego zaufania do dobieranych danych uznawane są odchylenia, jakie generują i dane i metody ich gromadzenia (por. Glaser, Strauss 2009, s. 58).

Istnieje jednak możliwość minimalizowania błędów pokrycia i podniesienia wiarygodności doboru realizowanego zgodnie z ideą doboru teoretycznego i nasycenia kategorii. Można bowiem zwiększyć zawartość informacyjną danych i jednocześnie minimalizować ryzyko wtórnego dublowania informacji oraz gromadzenia danych teoretycznie nieistotnych. Przed tym ryzykiem mocno przestrzegali Anselm Glaser i Barney Strauss, podkreślając przy tej okazji znaczenie niemarnowania czasu (2009, ss. 59, 61–62). Wspomniana możliwość mogłaby polegać na traktowaniu obszaru badań jako dziedziny, w której rozsiane są przypadki w niewiadomych miejscach, a następnie zastosowaniu postępowania dostosowanego do takich warunków. Ulegałaby wówczas redukcji rozbieżność właściwości dobranej próby i danych, jakich ona dostarcza, z właściwościami jednostek i stanem rzeczy cechującym całą zbiorowość. Sposoby radzenia sobie z takimi niewiadomymi w polu określonej dziedziny opracowane zostały w obszarze metody reprezentacyjnej, działu statystyki skupionego na zagadnieniu doboru próby z populacji i sposobów charakteryzowania zwykle nieznanymi cechami tej populacji. Jedną z najprostszych wśród możliwych propozycji jest wykorzystanie już na etapie wstępnego doboru jednostek rozwiązania zwanego doбором kwotowym. To jednak wymaga przynajmniej minimalnej wiedzy o zbiorowości. Poza tym w jego przypadku trzeba mieć pewność, że cechy, ze względu na które budowane będą kwoty, mają znaczenie dla celu badania. W sytuacji braku takiej wiedzy użyteczniejszy może więc okazać się dobór losowy. Różniące się między sobą przypadki oraz ich atrybuty rozproszone są w obrębie dziedziny, a nawet jej podzbiorów, w sposób trudny do

przewidzenia. Możemy jednak założyć, że możliwość trafienia na większość z nich zwiększy się, gdy każdemu przypadkowi lub atrybutowi zapewnimy takie same lub bardzo zbliżone szanse trafienia do próby. Spełnienie tego rodzaju warunku zachodzi właśnie przy zastosowaniu doboru losowego. Dobór ten nie powoduje ingerencji w strukturę i właściwości zbiorowości. Należy wyraźnie podkreślić, że jednostką losowania nie muszą też być poszczególne przypadki. Trudno bowiem oczekiwać, by sporządzenie operatu losowania (opracowana lista wszystkich jednostek danej kategorii) opierającego się na poszczególnych jednostkach było zawsze możliwe. Operat może składać się z jednostek będących w rzeczywistości zespołami poszczególnych przypadków (dowolnych *social units*), jak szkoły lub sekcje instytucji danego rodzaju. Można też losowaniu poddać miejsce w obrębie dziedziny (np.: ulica lub dzielnica miasta), z którego następować będzie systematyczny dobór samych jednostek, np. co trzeci napotkany przypadek. Nie wszystkie też jednostki w populacji mają taką samą wagę lub – mówiąc inaczej – dostarczać mogą równie bogatych informacji. W związku z tym możliwe jest zastosowanie warstwowania zbiorowości, tak by tym trafniej dobrać próbę. Polegałoby to na wcześniejszym rozeznaniu i zaplanowaniu podzbiorowości jednostek różniących się np. zakresem, złożonością i rodzajem informacji, jakie można uzyskać podczas gromadzenia danych. Warto zauważyć, że losowanie zespołowe oraz warstwowanie zbiorowości w pewnym zakresie zbiegają się z ideą dobierania grup porównawczych w ramach doboru teoretycznego (por. Glaser, Strauss, 2009). Polegają bowiem na wyodrębnianiu podzbiorowości według założonej zasady i dobieraniu z ich zakresu jednostek w sposób wyczerpujący albo częściowy. Tą zasadą mogłoby być kryterium teoretycznego znaczenia lub kryterium obszaru rzeczowego (Glaser, Strauss, 2009, s. 44, 46).

Podsumowując, teoria wyłaniająca się z danych mogłaby dostarczać warunków nakładanych na kolejne doборы prowadzone w oparciu o metodę reprezentacyjną. Przywołane propozycje mogą sprzyjać tworzeniu okoliczności minimalizujących ryzyko wystąpienia sytuacji zobrazowanej na rys. 1, a tym samym zwiększać szanse na osiągnięcie celów podkreślanych w kontekście nasylenia teoretycznego. Chodzi o rozszerzenie perspektywy możliwych „kombinacji empirycznych granic w danych” oraz kontrolowanie ogólności pojęciowej w następstwie docierania do wielu różnych grup jednostek. Jest to dyktowane wymaganiami we wstępnej fazie badań możliwie najpełniejszego pokrycia ogólnej charakterystyki zbioru wytypowanego celem badawczym oraz tworzenia warunków maksymalizujących wgląd w różnice tudzież podobieństwa jednostek i danych, których one dostarczają (Glaser, Strauss, 2009, s. 46, 48,

53, 59). Zaprezentowane przykłady nie wyczerpują jednak możliwości dopasowania schematu doboru próby do potrzeb badacza i okoliczności badania. Istnieją w tym zakresie liczne rozwiązania zwiększające poziom reprezentatywności próby, a tym samym informacyjnej zawartości danych empirycznych (Steczkowski, 1995; Lohr, 2010). Podstawowym czynnikiem ograniczającym stopień i zakres ich zastosowania jest głównie wyobraźnia badacza. Choć chcę też mocno podkreślić, że przykłady przywołałem, nie kierując się intencją absolutyzowania, lecz zdecydowanie z zamiarem wskazania alternatywnego sposobu maksymalizowania wiarygodności danych.

Zakończenie

Poszukiwanie odpowiedzi na pytanie o rozmiar próby nie prowadzi do bezpośrednich rozwiązań na wzór tych możliwych do uzyskania w ramach badań statystycznych. Stwarza raczej okazje do zweryfikowania samego pytania. Okazuje się bowiem, że to nie rozmiar próby ma kardynalne znaczenie, lecz sposób, w jaki jest dobierana. Reprezentatywność, możliwość tworzenia uogólnień czy też sama wiarygodność rezultatów badawczych uzależnione są od sposobu, w jaki próba została wyselekcjonowana, a nie bezpośrednio od jej liczebności. Jednakże mimo istniejących założeń organizacji próby w badaniach jakościowych oraz wyników badań nad najczęściej przyjmowanymi rozwiązaniami w zakresie jej rozmiaru, nie można zagwarantować, że oddziaływanie zniekształceń, wynikających z niedostatecznej reprezentacji w próbie, na poznanie zjawiska lub zbiorowości będzie mogło być uwidaczniane i minimalizowane.

Warto podkreślić to, o czym mowa była już wcześniej, że dokonując wyborów, badacz zdeterminowany jest przez właściwości, jakimi sam się cechuje (posiadana wiedza, cechy osobowości, właściwości percepcyjne, indywidualne doświadczenie itp.) oraz jakimi charakteryzuje się jego otoczenie (struktura środowiska, jego wyrazistość oraz dostępność itp.). Może to oznaczać, że okoliczności, w których przyszło mu dobierać próby, ograniczają możliwość dostarczenia do istotnych informacji o badanej zbiorowości lub zjawisku. Mając tego świadomość, łatwiej zabiegać o podnoszenie jakości materiału empirycznego. Dobór teoretyczny oraz początkowy dobór go poprzedzający nie muszą być obciążane mankamentami arbitralności.

Świadomość tego zdeterminowania nabiera jeszcze większego znaczenia w kontekście praktyki tworzenia uogólnień na podstawie rezultatów badawczych uzyskiwanych z próby. Próbuując tworzyć uogólnienia, wkraczamy na teren, w którym zastosowanie znajduje statystyka. Każde uogólnienie polega

bowiem na orzekaniu o wszystkich przypadkach danej kategorii. Mówiąc inaczej, skoro jakieś zjawisko jest przedmiotem badania i jego poznanie uzależnione jest od obserwacji poszczególnych jednostek jakoś uczestniczących w tym zjawisku lub z nim związanych, to problem ilości zawsze będzie istniał. Z tym problemem można bez kompleksów mierzyć się w obrębie projektów szeroko rozumianych badań jakościowych. Proces poznania zjawiska nie musi więc być jedynie oczekiwany lub zdany na okazję i niekontrolowane ryzyko niechcianej stronniczości w doborze środków. Istnieje możliwość szybszego docierania do zdarzeń, których nie dostrzeżemy albo ujrzymy znacznie później, jeśli poprzestaniemy na strategii brania tego, co jest w zasięgu ręki, i zdawania się głównie na własne intuicje. Wartościowe więc dla „badaczy jakościowych” może się okazać zawieszenie negacji osiągnięć metody reprezentacyjnej, która od początku wycelowana jest w szukanie możliwości i wypracowywanie procedur wiarygodnego uogólniania wniosków. Dobór próby wykorzystujący ideę szans wyboru jednostek ze zbiorowości generalnej stwarza warunki podnoszenia jakości uogólnień tworzonych w ramach badań jakościowych, i to względnie niezależnie od rozmiaru samej próby. Tego rodzaju rozwiązanie daje się umieścić w horyzoncie triangulacji metodologicznej, i choćby z tego powodu możliwe jest do pogodzenia z filozofią badań jakościowych.

Bibliografia

- Adler P., Adler P. (2012), *Expert Voices*, w: S. Baker, R. Edwards (red.), *How many qualitative interviews is enough? Discussion Paper* (s. 8–11), NCRM, ESRC. Pobrano 2014.12.12 z: <http://eprints.ncrm.ac.uk/2273/>.
- Babbie E. (2006), *Badania społeczne w praktyce*, przeł. W. Betkiewicz, M. Bucholc, P. Gadowski, J. Haman, A. Jasiewicz-Betkiewicz, A. Kloskowska-Dudzińska, M. Kowalski, M. Mozga, PWN, Warszawa.
- Baker S., Edwards R. (2012), *How many qualitative interviews is enough. Discussion Paper*, NCRM, ESRC. Pobrano 2014.12.12 z: <http://eprints.ncrm.ac.uk/2273/>.
- Bowen G. (2008), *Naturalistic inquiry and the saturation concept: a research note*, “Qualitative Research”, 8 (1), 137–142, DOI: <http://dx.doi.org/10.1177/1468794107085301>.
- Charmaz K. (2009), *Teoria ugruntowana. Praktyczny przewodnik o analizie jakościowej*, przeł. B. Komorowska, PWN, Warszawa.
- Cohen L., Manion L., Morison K. (2007), *Research Method in Education*, Routledge, London–New York.

- Creswell J. (2007), *Qualitative Inquiry & Research Design. Choosing Among Five Approaches*, Sage Publications, Thousand Oaks–London–New Delhi.
- Flick U. (2011a), *Jakość w badaniach jakościowych*, przeł. P. Tomanek, PWN, Warszawa.
- Flick U. (2011b), *Projektownie badań jakościowych*, przeł. P. Tomanek, PWN, Warszawa.
- Gibbs G. (2011), *Analizowanie danych jakościowych*, przeł. M. Brzozowska-Brywczyńska, PWN, Warszawa.
- Glaser B. G., Strauss A. L. (2009), *Odkrywanie teorii ugruntowanej. Strategie badania jakościowego*, przeł. M. Gorzko, Zakład Wydawniczy NOMOS, Kraków.
- Gorzko M. (2001), *Problem reprezentatywności próby w zastosowaniach metody biograficznej*, „Firma i Rynek”, 20 (3), 68–71.
- Guba E. (1981), *Criteria for Assessing the Trustworthiness of Naturalistic*, “Educational Communication and Technology”, 29 (2), 75–91.
- Guest G., Bunce A., Johnson L. (2006), *How Many Interviews Are Enough? An Experiment with Data Saturation and Variability*, “Field Methods”, 18 (1), 59–82.
- Hajduk Z. (2011), *Ogólna metodologia nauk*, Wydawnictwo KUL, Lublin.
- Hall T., Herron T., Pierce B. (2006), *How Reliable Is Haphazard Sampling?*, “The CPA Journal”, 76 (1), 26–27.
- Kamiński S. (1989), *Jak filozofować? Studia z metodologii klasycznej*, Towarzystwo Naukowe KUL, Lublin.
- Kołodziej A. (2013), *Teoria estymacji w praktyce badań społecznych*, Difin, Warszawa.
- Konecki K. (2000), *Studia z metodologii badań jakościowych. Teoria ugruntowana*, PWN, Warszawa.
- Lewis J., Ritchie J. (2003), *Generalising from Qualitative Research*, w: J. Ritchie, J. Lewis, (red.), *Qualitative research practice. A guide for social science students and researchers*, SAGE Publication, Thousand Oaks–New Delhi–Londyn, s. 263–286.
- Lohr S. (2010), *Sampling: Design and Analysis*, Brooks/Cole, Boston.
- Marshall B., Cardon P., Poddar A., Fontenot R. (2013), *Does sample size matter in qualitative research?: A review of qualitative interviews in is research*, “Journal of Computer Information System”, 54 (1), 11–22.
- Mason M. (2010), *Sample Size and Saturation in PhD Studies Using Qualitative*, “Forum Qualitative Sozialforschung”, 11 (3) / “Forum: Qualitative Social Research”, Art. 8. Pobrano 2014.06.12 z: <http://nbn-resolving.de/urn:nbn:de:0114-fqs100387>.
- Morse J., Barret M., Mayan M., Olson K., Spiers J. (2002), *Verification Strategies for Es-*

Establishing Reliability and Validity in Qualitative Research, "International Journal of Qualitative Methods", 1(2), 13–22.

Nowak S. (2007), *Metodologia badań społecznych*, PWN, Warszawa.

O'Reilly M., Parker N. (2013), 'Unsatisfactory Saturation': a critical exploration of the notion of saturated sample sizes in qualitative research, "Qualitative Research", 13 (2), 190–197.

Pasikowski S. (2014a), *Kultura metodologiczna i raportowanie badań empirycznych publikowanych w wiodących czasopismach poświęconych zagadnieniom edukacji*, "Kultura i Edukacja", 2, 103–133.

Pasikowski S. (2014b), *Transparency of research published in the leading polish educational journals*, w: A. L. L. Gómez Chova (red.), *INTED 2014 Proceedings*, IATED Academy, Valencia, s. 4152–4159.

Ritchie J., Lewis J., Elam G. (2003), *Designing and Selecting Samples*, w: J. Ritchie, J. Lewis (red.), *Qualitative Research Practice. A Guide for Social Science Students and Researchers*, SAGE Publications, Thousand Oaks–New Delhi–London, s. 77–108.

Sandelowski M. (1993), *Rigor or rigor mortis: The problem of rigor in qualitative research revisited*, "Advances in Nursing Science", 16 (2), 1–8, DOI: <http://dx.doi.org/10.1097/00012272-199312000-00002>.

Siekierski K. (2002), *Mit próby reprezentatywnej*, „Marketing w Praktyce”, 1, 10–12.

Silverman D. (2008a), *Interpretacja danych jakościowych*, przeł. M. Głowacka-Grajper, J. Ostrowska, PWN, Warszawa.

Silverman D. (2008b), *Prowadzenie badań jakościowych*, przeł. J. Ostrowska, PWN, Warszawa.

Steczkowski J. (1995), *Metoda reprezentacyjna w badaniach zjawisk ekonomiczno-społecznych*, PWN, Warszawa–Kraków.

Szreder M. (2010), *Metody i techniki sondażowych badań opinii*, Polskie Wydawnictwo Ekonomiczne, Warszawa.

Taylor J. (1995), *Wstęp do rachunku błędu pomiarowego*, przeł. A. Bobiński, R. Bożek, PWN, Warszawa.

Thomson S. B. (2011), *Sample Size and Grounded Theory*, "Journal of Administration & Governace", 5 (1), 45–52.

* W prezentowanym tekście wykorzystane zostały wybrane idee zawarte w:

Pasikowski S. (2015), *Granice teoretycznego nasycenia*, "Rocznik Lubuski", nr 41 (1), 33-40.

