

Andrzej Siemaszko

METODOLOGICZNE PROBLEMY BADAN TYPU *SELF-REPORT*

I. WSTĘP

Często przyjmowany schemat badań w kryminologii wygląda następująco: do grupy eksperymentalnej dobiera się osoby, które weszły w kolizję z prawem (były podejrzane, sądzone lub skazane za popełnienie przestępstwa), do grupy kontrolnej zaś — tzw. nieprzestępców, a więc osoby, co do których zakłada się, że nie popełniły przestępstwa. Ten schemat analizy przedstawiliśmy jednak w sposób nieco wyidealizowany. W rzeczywistości bowiem w wielu badaniach nie stosuje się w ogóle grupy kontrolnej, grupy eksperymentalne zaś stanowią jeszcze bardziej wyselekcjonowane kategorie przestępców — więźniów lub wychowanków zakładów poprawczych. W każdym razie jednak przedmiotem badania czyni się często osoby, którymi zainteresował się już aparat wymiaru sprawiedliwości.

Zakłada się więc z jednej strony — co zrozumiałe — że osoby stanowiące grupę eksperymentalną na pewno są przestępcami, z drugiej zaś — że osoby, które dobiera się do grupy kontrolnej, z pewnością przestępstw nie popełniają. To drugie założenie jest już jednak mocno wątpliwe. Wiadomo bowiem od lat, że także i wśród tzw. nieprzestępców niemały odsetek stanowią osoby, które również dokonują przestępstw, choć przestępstwa te nie zostają ujawnione.

Mimo tej zasadniczej wady odchodzenie od przedstawionego na wstępie wadliwego schematu badawczego odbywa się jednak dość opieszale. A przecież tak projektowane zamierzenia badawcze mogłyby być zasadne wówczas jedynie, gdyby uznawać fakt przebywania w zakładzie za wyłączne kryterium przestępczości. Jest to wszakże założenie nieuzasadnione, szczególnie w badaniach o charakterze etiologicznym. Jeżeli bowiem interesuje nas, co sprawia, w największym uproszczeniu, że Iksiński popełnia przestępstwa, to jeśli decydujemy się uznać, że dzieje się tak dlatego, że brak mu czegoś, co posiada Igrekowski, który przestrzegą prawa, lub też że Igrekowski ma coś, czego brak Iksińskiemu-przestępcy, musimy naj-

pierw być pewni, że Igrekowski z pewnością nie narusza przepisów kodeksu karnego. W przeciwnym bowiem razie dalszy ciąg rozumowania wynikającego z kanonu jedynej różnicy staje się chybiony.

Badanie osób, które weszły w orbitę zainteresowań instytucji formalnej kontroli społecznej ma także i inne wady. Przede wszystkim zaliczyć do nich należy brak reprezentatywności prób dobieranych na podstawie tego kryterium, co wynika z dużej selektywności działania aparatu ścigania i wymiaru sprawiedliwości. Wspominaliśmy już, że nie każde przestępstwo zostaje wykryte, lecz także nie każdy zatrzymany przestępca dostaje się w tryby maszyny wymiaru sprawiedliwości (zostaje np. aresztowany, osądzony i skazany), a jeżeli nawet, to nie przechodzi przez wszystkie jej tryby. Z rozmaitych przyczyn postępowanie karne kończy się wobec nich na etapie postępowania przygotowawczego (np. umorzenie lub warunkowe umorzenie) lub na etapie postępowania sądowego (np. umorzenie, warunkowe umorzenie, warunkowe zawieszenie wykonania kary czy uniewinnienie). Z wielu względów może nawet nie dojść do wykonania bezwzględnej kary pozbawienia wolności. Widzimy zatem, jak w kolejnych stadiach postępowania karnego maleje i — co ważniejsze — selekcjonuje się według trudnych do ustalenia kryteriów przysła próba „przestępców”.

Charakter opracowania uniemożliwia analizowanie wszystkich okoliczności warunkujących selektywność działania wymiaru sprawiedliwości. O jednej należy však wspomnieć. Stanowi ją pozycja społeczno-ekonomiczna sprawcy przestępstwa. Statystyki sądowe wskazują jednoznacznie, że wśród skazanych dominują osoby wywodzące się z niższych warstw społeczno-ekonomicznych. Wynika to z wielu przyczyn, których nie będziemy w tym miejscu zgłębiać. Ważne jest bowiem jedno. O ile uznamy ten fakt za bezsporny (a potwierdzają go niezliczone badania kryminologiczne), to warto zdawać sobie sprawę, że wszystkie zmienne, które zwykle się traktować jako „przyczyny przestępczości”, mogą być w istocie skorelowane nie tyle z przestępczością, co z niską pozycją społeczno-ekonomiczną jako taką. A zatem te zmienne, które nierzadko bezrefleksyjnie uważa się za „etiologiczne czynniki przestępczości”, mogą stanowić co najwyżej „etiologiczne czynniki przestępczości niższych warstw społeczno-ekonomicznych” lub w przypadkach skrajnych po prostu cechy charakterystyczne tych warstw, cechy, które w ogóle nie mają związku z przestępczością. Oczywiście są to truizmy metodologiczne. Jednak te oczywiste prawdy z trudem bywają akceptowane przez wielu badaczy zjawisk patologii społecznej.

Sygnalizowana uprzednio selektywność działania wymiaru sprawiedliwości rysuje się, jak się wydaje, szczególnie ostro w przypadku przestępczości nieletnich. Wynika bowiem nie tylko z praktyki, ale wręcz z rozwiązań ustawowych. Sędziom dla nieletnich pozostawia się zazwyczaj swobodę tak w ocenie konkretnego czynu, jak i w sferze wyboru ewentualnych sankcji. Nie inaczej działa sądownictwo dla nieletnich także i w naszym kra-

ju. Oczywiście, jest to rozwiązanie trafne. Warto jednak zdawać sobie sprawę ze skutków, jakie dla badań kryminologicznych wynikają z faktu, że sędzia dla nieletnich ma dużą władzę dyskrecyjną. Posłużmy się przykładem. Otóż jeżeli przed obliczem sędziego dla nieletnich staje dwóch chłopców oskarżonych o ten sam rodzaj kradzieży, z tym że pierwszy z nich pochodzi z tzw. dobrego domu, jego rodzice są dobrze sytuowani i mogą mu zapewnić właściwą opiekę, drugi zaś wywodzi się ze środowiska tzw. marginesu społecznego, jego rodzice są po rozwodzie, a na dodatek każde z nich nadużywa alkoholu, to nietrudno zgadnąć, który z chłopców z naszego przykładu trafi, co najwyżej, pod dozór kuratora, a wobec którego sprawę przekaże się szkole lub umorzy. Potem okazuje się z badań, że „nieletni przestępcy” wywodzą się ze środowiska marginesu społecznego, z rozbitych domów i z rodzin, w których nadużywa się alkoholu.

Badania inkarcerowanych przestępców rodzą jeszcze inne problemy metodologiczne, których już to się nie dostrzega, już to ignoruje. Przestępcy znajdujący się w więzieniu czy zakładzie poprawczym mają już za sobą z reguły wiele kontaktów z reprezentantami formalnej i nieformalnej kontroli społecznej. Inkarceracja stanowi bowiem najczęściej zakończenie długiego łańcucha interakcji przestępcy — tak z najbliższym otoczeniem, jak i z przedstawicielami instytucji formalnej kontroli społecznej. Nie pozostaje to oczywiście bez wpływu na jego osobowość, sposób postrzegania siebie, innych ludzi i świata. Jeżeli inni nastawieni są do nas wrogo, my odpłacamy im tym samym. Jeżeli nas podejrzewają, my w końcu sami stajemy się podejrzliwi. Jeżeli nas odrzucają, my też mamy tendencję do odrzucania innych. Tak więc przestępca trafiający do więzienia lub zakładu poprawczego jest już nierzadko innym człowiekiem, niż był uprzednio. A zatem prowadząc etiologiczne badania przestępczości na osobach inkarcerowanych bardzo łatwo można uznać skutki za przyczyny. Nie trzeba zaś szerzej uzasadniać tezy, że warunki więzienne pogłębiają z reguły proces transformacji osobowości, prowadząc często nawet do poważnych zaburzeń psychicznych. Jeżeli więc w rezultacie badania więźniów stwierdzi się przykładowo, że mają oni nieprawidłowe profile w teście MMPI, to na takiej podstawie można powiedzieć co najwyżej, że dla osób przebywających w więzieniu charakterystyczny jest taki właśnie profil wyników testu MMPI. Tymczasem tego rodzaju ustalenia bywają traktowane jako czynniki doprowadzające do przestępczości generalnie.

Biorąc powyższe pod uwagę, w światowej kryminologii i socjologii dewiacji rezygnuje się stopniowo w badaniach etiologicznych ze schematu badawczego, analizowanego powyżej, na rzecz badania przestępczości w populacjach nie zinstytucjonalizowanych, a nawet w populacji generalnej. Wśród tego typu badań na uwagę zasługują szczególnie badania typu *self-report*. Na czym najogólniej polega ten rodzaj badań? Otóż przede wszystkim odmienne jest kryterium przestępczości czy szerzej — dewiacji. W ba-

daniach typu *self-report* nieistotny jest fakt zatrzymania, aresztowania, oskarżenia, osądzenia czy skazania sprawcy przestępstwa. W badaniach tych przestępcą jest ten, kto sam się przyznaje do popełnienia przestępstwa, bez względu na to, czy przestępstwo to zostało ujawnione, czy też nie. W metodzie *self-report* respondenci otrzymują na ogół do wypełnienia anonimowy kwestionariusz, który obok innych pytań zawiera listę zachowań przestępczych lub (i) dewiacyjnych. Zadaniem badanych jest udzielenie odpowiedzi na pytanie o to, czy i ewentualnie jak często dopuszczali się takich czynów. Uzyskane w ten sposób wyniki bywają najczęściej wykorzystywane dwojako.

Po pierwsze, za pomocą badania typu *self-report* można starać się oszacować rozmiary tzw. ciemnej liczby przestępstw w populacji generalnej. Nie jest to jednak łatwe. Nie wiadomo bowiem przede wszystkim, czy czyny, do których dokonania przyznają się badani, byłyby uznane za przestępstwa przez organy ścigania i wymiaru sprawiedliwości w przypadku ich ujawnienia. Na przeszkodzie ewentualnym szacunkom ciemnej liczby przestępstw stoi ponadto najczęściej charakter próby, ograniczający w poważnym stopniu zakres dopuszczalnych uogólnień. Przykładowo wymienione trudności, a jest ich w rzeczywistości znacznie więcej, sprawiają, że na podstawie wyników badań typu *self-report* można dokonywać jedynie przybliżonych i mało precyzyjnych szacunków rozmiarów faktycznej przestępczości.

Po drugie, rezultaty uzyskiwane przez respondentów w badaniu typu *self-report* służyć mogą do tworzenia zmiennych niezależnych „zachowanie przestępcze” czy „zachowanie dewiacyjne”, które są podstawą dalszych analiz. W procedurze tej dokonuje się zazwyczaj podziału badanej zbiorowości na „przestępców” i „nieprzestępców” lub („dewiantów” i „niedewiantów”) właśnie na podstawie rodzajów ich odpowiedzi na pytania o zachowania przestępcze (dewiacyjne). Następnie dokonuje się dalszych analiz (np. przyczynowych) z uwzględnieniem określonych zmiennych niezależnych.

Poczynając od lat sześćdziesiątych metoda *self-report* staje się systematycznie normalnym, rutynowym sposobem pomiaru rozmiarów, struktury i uwarunkowań przestępczości i innych zachowań dewiacyjnych, wypierając stopniowo tradycyjne badania osób pozbawionych wolności. W niektórych krajach (np. w Stanach Zjednoczonych i w Kanadzie) obserwuje się obecnie nawet wyraźne zachwianie proporcji między badaniami typu *self-report* a badaniami, które umownie można określić jako tradycyjne: te drugie spotyka się już bardzo rzadko; szczególnie dotyczy to badań przestępczości nieletnich. Tymczasem w Polsce metoda *self-report* jest praktycznie nieznaną i niezmiernie rzadko stosowaną. Jednym z celów niniejszej pracy jest zatem przybliżenie polskiemu czytelnikowi-specjaliście od zagadnień dewiacji, patologii społecznej czy przestępczości tej interesującej

i wartościowej poznawczo metody. Być może, dzięki temu badania typu *self-report* będą w przyszłości prowadzone na szerszą niż dotychczas skalę, bowiem ich zalety, jak się wydaje, są niewątpliwe.

Badanie typu *self-report* to zasadniczo badanie za pomocą anonimowej ankiety audytoryjnej. Dlatego też wydaje się, iż metodologiczne problemy, które wiążą się z zastosowaniem tej techniki prowadzenia badania, nie są specyficzne wyłącznie dla niej. Większość zagadnień metodologicznych dotyczy w istocie w tym samym stopniu badań typu *self-report*, co wszystkich innych badań za pomocą ankiety anonimowej. Przyjmuje się jednak na ogół, że w badaniach typu *self-report* stykamy się z problemami i dylematami, które nie występują w innych badaniach z użyciem anonimowej ankiety audytoryjnej. Szczególnie ostro rysuje się, przykładowo, kwestia szczerości i prawdomówności respondentów. W badaniach, o których mowa, pytamy przecież o sprawy drażliwe. Przyznanie się do tego, że ukradło się kiedyś określoną sumę pieniędzy czy też że dokonano się włamania lub pobicia, jest w istocie przyznaniem się do faktu popełnienia przestępstwa. A czy łatwo jest się do tego przyznać, szczególnie gdy nie ma się całkowitej pewności, czy aby wypełniona ankieta służyć będzie wyłącznie celom naukowym? Obawy tego rodzaju mogą być szczególnie częste w tych krajach, w których ankieta socjologiczna stanowi w dalszym ciągu pewne *novum* dla przeciętnego obywatela. Zarówno kwestia prawdomówności respondentów, która dla badań typu *self-report* wydaje się kluczowa, jak też i inne zagadnienia metodologiczne, specyficzne dla tej metody prowadzenia badań, sprawiły, że istnieje już dość pokaźna literatura poświęcona głównie lub wyłącznie metodologicznym problemom badań typu *self-report*². Niezależnie od tego większość badaczy stosujących metodę *self-re-*

¹ Choć istnieją też i inne sposoby uzyskiwania informacji o nie ujawnionych zachowaniach dewiacyjnych, o czym szerzej w rozdziale III.

² Poniżej przytaczamy ważniejsze pozycje bibliograficzne traktujące o metodologii badań *self-report*. Bibliografia ułożona jest w porządku chronologicznym, z tym że jeżeli kilka prac tego samego autora pochodzi z różnych okresów, cytowane są one w jednym miejscu, poczynając od tekstu najwcześniejszego. Zob. min. I. Nye, J. F. Short: *Scaling Delinquent Behavior*, „American Sociological Review”, 1957, vol. 22; J. F. Short, I. Nye: *Reporter Behavior as a Criterion of Deviant Behavior*, „Social Problems”, 1957, vol. 5; J. F. Short, I. Nye: *Extent of Unrecorded Juvenile Delinquency: Tentative Conclusions*, „The Journal of Criminal Law, Criminology and Police Science”, 1958, vol. 49; I. Nye: *Family Relationships and Delinquency*, New York 1958; M. Gold: *Status Forces in Delinquent Boys*, Ann Arbor 1963; M. Gold: *Undetected Delinquent Behavior* „Journal of Research in Crime and Delinquency”, 1966, vol. 3; M. Gold: *Deviant Behavior in an American City*, Belmont 1970; R. Dentier, L. Monroe: *Early Adolescent Theft*, „American Sociological Review”, 1961, vol. 26; I. Deutscher: *Ords and Deeds: Social Science and Social Policy*, „Social Problems”, 1966, vol. 13;

T. Hirschi: *Causes of Delinquency*, Berkeley 1969; N. Heise: *Norms and Individual Patterns in Students Deviance*, „Social Problems”, 1968, vol. 16; R. H a r d t,

port także sygnalizuje pewne kwestie metodologiczne przy okazji referowania uzyskanych rezultatów³. Tak więc materiał dotyczący metodologicznej strony badań typu *self-report* jest pokaźny. I na nim właśnie, a nie na bardziej generalnych problemach metodologicznych techniki ankietowej, skoncentrujemy uwagę w niniejszym opracowaniu.

Układ pracy przedstawia się następująco. W rozdziale I przedstawimy krótki rys historyczny badań metodą *self-report*, koncentrując się szcze-

S. Peterson: *Arrests of Self and Friends as Indicators of Delinquent Involvement*, „Journal of Research in Crime and Delinquency”, 1968, vol. 5; R. Hardt, S. Peterson-Hardt: *On Determining the Quality of Delinquency Self-report Method*, „Journal of Research in Crime and Delinquency”, 1977, vol. 14; L. Gould: *Who Defines Delinquency: a Comparison of Self-reported and Officially Reported Indices for Racial Groups*, „Social Problems”, 1969, vol. 16; L. Gould: *Juvenile Entrepreneurs*, „American Journal of Sociology”, 1969, vol. 78; N. Christie, J. Andeneas, S. Skirbekk: *A Study of Self-reported Crime*, „Scandinavian Studies in Criminology”, 1965, vol. 94; R. Hardt, G. Bodine: *Development of Self-report Instruments in Delinquency Research: a Conference Report*, Syracuse 1965; J. Hackler, B. Lautt: *Systematic Bias in Measuring Self-reported Delinquency*, „Canadian Review of Sociology and Anthropology”, 1969, vol. 92; M. Hindelang: *The Social Versus Solitary Nature of Delinquent Involvement*, „The British Journal of Criminology”, 1971, vol. 11; M. Hindelang: *With a Little Help from Their Friends: Group Participation in Reported Delinquent Behavior*, „The British Journal of Criminology”, 1976, vol. 16; J. Clark, L. Tift: *Polygraph and Interview Validation of Self-reported Behavior*, „American Sociological Review”, 1966, vol. 31; W. Belson: *The Extent of Stealing by London Boys and Some of Its Origins*, London 1968; W. Belson: *Juvenile Theft: The Causal Factors*, London 1975; Arnold: *Continuities in Research: Scaling Delinquent Behavior*, „Social Problems”, 1965, vol. 64; Kulik, Stein, Sabrin: *Disclosure of Delinquent Behavior under Conditions of Annonimity and Nonanonimity*, „Journal of Consulting and Clinical Psychology”, 1968, vol. 32; W. Bowers: *Questions Sequencing Effects on Response to Sensitive Questions in Self-administered Questionnaire*, Steak Research Center 1971; D. Farrington: *Self-reports of Deviant Behavior: Predictive and Stable?*, „Journal of Criminal Law and Criminology”, 1973, vol. 64; M. Krohn, G. Waldo, T. Chiricos: *Self-reported Delinquency: a Comparison of Structured Interview and Self-administered Checklist*, „The Journal of Criminal Law and Criminology”, 1974, vol. 65; J. Blackmore: *The Relationships between Self-reported Delinquency and Official Convictions among Adolescent Boys*, „The British Journal of Criminology”, 1974, vol. 14; A. Liska: *Emergent Issues in the Attitude-behavior Consistency Controversy*, „American Sociological Review”, 1974, vol. 39; R. Smart: *Recent Studies of the Validity and Reliability of Self-reported Drug Use*, „Canadian Journal of Criminology and Corrections”, 1975, vol. 17; M. Hindelang, T. Hirschi, J. Weis: *Correlates of Delinquency: the Illusion of Discrepancy between Self-report and Official Measures*, „American Sociological Review”, 1979, vol. 44; D. Elliott, S. Ageton: *Reconciling Differences in Eestimates of Delinquency*, „American Sociological Review”, 1980, vol. 45; S. Cordrey, K. Polk: *The Implications of Respondent Loss in Panel Studies of Deviant Behavior*, „Journal of Research in Crime and Delinquency”, 1983, vol. 20.

³Np. Axenroth: *Social Class and Delinquency*, „Journal of Research in Crime and Delinquency”, vol. 20, 1983.

golnie na prezentacji badań Shorta i Nye'a, od których datuje się trwająca do dziś popularność badań tego rodzaju. W rozdziale tym omówimy ponadto badania typu *self-report* prowadzone w Polsce. W rozdziale II zanalizujemy najważniejsze problemy metodologiczne, jakie nastroczą badania typu *self-report* oraz rozważymy kwestię, na ile jest to metoda godna zaufania, a także które z metodologicznych zagadnień referowanych uprzednio wydają się być już empirycznie rozstrzygnięte, które zaś wymagają dalszych badań.

II. PRZEGLĄD BADAŃ PROWADZONYCH METODĄ *SELF-REPORT*

1. UWAGI WPROWADZAJĄCE

Zanalizujemy poniżej wybrane wyniki badań poświęconych kwestii nie ujawnionych zachowań dewiacyjnych, w których stosowano metodę *self-report*. Badań tego rodzaju prowadzi się obecnie bardzo wiele, ponieważ *self-report* stał się już powszechnie stosowanym narzędziem pomiaru zachowań dewiacyjnych. Dlatego też nie sposób omówić dokładniej nawet niewielką część badań tego rodzaju. Należało więc dokonać ostrej selekcji w bardzo bogatym materiale empirycznym. Do prezentacji wybraliśmy badania, z którymi zazwyczaj wiąże się początek stosowania metody *self-report* w kryminologii i socjologii dewiacji. W szczególności zajmiemy się metodologiczną i merytoryczną stroną serii badań przeprowadzonych przez Shorta i Nye'a, które przez wiele lat stanowiły wzorzec do prowadzenia innych prac tego rodzaju. W drugiej części rozdziału scharakteryzujemy próby wykorzystywania metody *self-report* w badaniach polskich.

2. POCZĄTKI BADAŃ TYPU *SELF-REPORT*

Historia badań typu *self-report* jest stosunkowo krótka, datuje się bowiem od 1946 r., kiedy to zostały opublikowane rewelacyjne jak na owe czasy wyniki badań Porterfielda⁴ oraz Murphy'ego i innych⁵. Rezultaty tych drugich badań omówimy w największym skrócie, ponieważ nie zastosowano w nich „klasycznego *self-report*”, tj. anonimowej ankiety, lecz badanie przypadków i wywiady indywidualne. Autorzy w ciągu badania trwającego pięć lat ustalali liczbę czynów przestępczych dokonywanych przez 114 chłopców w wieku 11—15 lat⁶. Okazało się, że jedynie 14 bada-

⁴ A. Porterfield: *Youth in Trouble*, Forth Worth 1946, Szersze omówienie badań Porterfielda można znaleźć w pracy: J. Jasiński: *Kształtowanie się przestępczości nieletnich w Polsce w latach 1951—1960 w świetle statystyki sądowej*, „Archiwum Kryminologii”, 1964, t. II.

⁵ F. Murphy, M. Shirley, H. Witmer: *The Incidence of Hidden Delinquency*, „American Journal of Orthopsychiatry”, 1946, vol. 16.

⁶ Należy zaznaczyć, że były to „przestępstwa nieletnich” zgodnie z ustawodawstwem stanu Massachusetts w USA, tak więc nie wszystkie z tych czynów byłyby uznane za przestępstwa, przykładowo, w Polsce.

nych nie dopuściło się w tym czasie czynu, który kwalifikowałby ich jako „nieletnich przestępców”. Pozostali dokonali blisko sześć i pół tysiąca takich czynów. Lecz mimo to jedynie w stosunku do niecałych stu czynów wszczęto postępowanie karne. Przytłaczająca większość pozostała nie ujawniona, w tym czyny, które powszechnie uważa się za świadczące o dużym stopniu demoralizacji sprawcy (kradzieże, ucieczki z domu itp.). Badani popełnili przeszło 600 czynów tego rodzaju, które byłyby uznane za przestępstwa wówczas, gdyby dopuściła się ich osoba dorosła, lecz tylko 11% tych czynów o wyraźnych znamionach przestępstwa stało się podstawą wszczęcia postępowania.

Prawozorem dla współczesnych badań *self-report* stało się jednak opublikowane w tym samym czasie badanie Porterfielda⁷, ponieważ po raz pierwszy zastosowano w nim anonimowy kwestionariusz.

Punktem wyjścia badania Porterfielda była analiza przestępstw dokonanych przez dwa tysiące nieletnich, którzy odpowiadali za swe czyny przed sądem. Autor podzielił te przestępstwa na przeszło pięćdziesiąt typów, poczynając od czynów tak błahych, jak zakłócenie spokoju publicznego czy wagarowanie, aż do bardzo poważnych, jak zabójstwo. Następnie ustalał za pomocą kwestionariusza zawierającego owe kategorie przestępstw, czy i kiedy dopuścili się ich studenci (437 osób obojga płci). Wśród badanych nie znalazł się nikt, kto nie popełniłby choć raz w swoim życiu przestępstwa, za które odpowiadali badani uprzednio nieletni. Na jednego studenta przypadało średnio 17,6 przestępstw w okresie poprzedzającym studia, w okresie zaś studiów — przeciętnie 11,6 przestępstw. Badane studentki dokonały przeciętnie 4,7 przestępstw w okresie poprzedzającym studia. Przykładowo: w grupie studentów 23% przyznało się do kradzieży drobnych przedmiotów, 10% do kradzieży ze sklepu, 20% badanych posiadało rzeczy pochodzące z kradzieży, 5,5% zaś przyznało się do usiłowania gwałtu. Odsetki studentów przyznających się do popełniania różnego rodzaju przestępstw w okresie studiów były nieco niższe, co jest zrozumiałe, jeśli zważyć, że pytania dotyczące okresu poprzedzającego studia obejmowały dłuższy przedział czasu, a zatem badani mieli wówczas więcej okazji, by dokonać któregoś z przestępstw, wyszczególnionych w kwestionariuszu Porterfielda. Zgodnie z oczekiwaniami rozmiary nie ujawnionej przestępczości studentek były znacznie mniejsze. Mówimy tu o przestępczości nie ujawnionej, ponieważ — jak ustalił Porterfield — jedynie kilka procent badanych studentów odpowiedziało przed sądem w związku z czynami, do których popełnienia się przyznali. Reszta przestępstw pozostała nie ujawniona.

Zarówno wyniki badań Murphy’ego, jak i Porterfielda wskazują, że gdyby określać jako „przestępcę” każdego, kto popełnia czyn sprzeczny

⁷ A. Porterfield: op. cit.

z prawem, to w obydwu populacjach osób, których większość nie miała do czynienia z organami ścigania w związku ze swym naruszeniem prawa, dominują w istocie „przestępcy”.

Duży rezonans uzyskało w społeczeństwie amerykańskim badanie Wallersteina i Wyle'a, którego rezultaty zostały opublikowane w rok później⁸. Także i w tym badaniu zastosowano anonimowy kwestionariusz, Ustalono w ten sposób rozmiary przestępczości blisko 1700 osób dorosłych obojga płci. Średnia liczba przestępstw popełnionych po ukończeniu 16 roku życia wynosiła u mężczyzn 18, u kobiet zaś 11. Badani przyznawali się do popełnienia mniejszej liczby przestępstw w okresie, kiedy byli młodzi: na mężczyznę przypadało średnio 3,2 przestępstw popełnionych przed ukończeniem szesnastu lat, na kobietę natomiast 1,6.

Porównanie wyników powyższych badań wskazuje na istnienie bardzo dużych rozbieżności, jeśli chodzi o odsetki respondentów, którzy przyznali się do popełnienia przestępstw przed ukończeniem lat szesnastu (badania Wallersteina i Wyle'a, Porterfielda, Murphy'ego i innych).

Niezależnie od odmiennej metodologii tych badań i różnego stopnia ich metodologicznego zaawansowania (a także odmiennego katalogu uwzględnianych czynów) sądzić można, iż w badaniu Wallersteina i Wyle'a wystąpiła naturalna tendencja dorosłych osób do zapominania dość odległych w czasie zdarzeń. Z tego m.in. względu większość kolejnych badań prowadzona była wśród młodzieży.

Seria doskonałych badań Shorta i Nye'a⁹ jest jak gdyby kłamrą spinającą historię badań *self-report* z terażniejszością. Przez wiele lat kwestionariusz i metodologia wypracowane przez tych autorów stanowiły wzorzec, wedle którego prowadzono inne badania tego rodzaju. Dlatego też wypadnie nieco szerzej zreferować metodologię tych badań, jak też niektóre ich rezultaty.

Short i Nye uzyskali drogą doboru losowego dwie próby uczniów klas dziewiątych—jedenastych szkół średnich (n=2350). Pierwszą z prób tworzyli uczniowie wylosowani do badania ze szkół w trzech niewielkich miastach (10—30 tys. mieszkańców). Próba druga obejmowała uczniów szkół średnich w terenów wiejskich lub podmiejskich z odmiennego geograficznie, ekonomicznie i kulturowo obszaru Stanów Zjednoczonych (n=596). Próbę trzecią stanowili chłopcy i dziewczęta z zakładu poprawczego (n = = 230). Tak więc łącznie przebadano przeszło 3 tys. dziewcząt i chłopców.

Short i Nye kierowali się następującymi kryteriami doboru zachowań do kwestionariusza *self-report*. Przede wszystkim uwzględniano takie czy-

⁸ J. Wallerstein, C. Wyle: *Our Law-abiding lawbreakers*, „Probation”, 1947, vol. 4. Szersze omówienie badań Wallersteina i Wyle'a można znaleźć w: P. Horoszowski: *Kryminologia*, Warszawa 1965.

⁹ J. Nye, J. F. Short: *op. cit.*; J. F. Short, J. Nye: *op. cit.* (1957, 1958); J. Nye: *op. cit.*

ny, które zgodnie z ustawodawstwem amerykańskim mogłyby być uznane za przestępstwa nieletnich. Spośród nich wybrano takie, które:

1) gwarantowały odpowiednio wysokie prawdopodobieństwo, że przyzna się do nich wystarczająco do celów analizy statystycznej odsetek badanych;

2) były popełniane przez poważny odłam społeczeństwa amerykańskiego;

3) zawierały w sobie różny stopień społecznego niebezpieczeństwa (tj. zarówno czyny błahe, jak i stosunkowo poważne).

Kwestionariusz liczył 23 pytania, z tym że nie wszystkie z nich zostały następnie uwzględnione przy konstruowaniu skal dewiacji. Kwestionariusz Shorta i Nye'a składał się z pytań o następujące zachowania:

- 1) jazda samochodem bez zezwolenia lub prawa jazdy;
- 2) opuszczenie szkoły bez usprawiedliwienia;
- 3) nieposłuszeństwo rodzicom;
- 4) bójka na pięści;
- 5) mówienie nieprawdy;
- 6) ucieczki z domu;
- 7) zawieszenie w prawach ucznia lub usunięcie ze szkoły;
- 8) jawne sprzeciwianie się woli rodziców;
- 9) przekraczanie szybkości lub nieostrożna jazda samochodem;
- 10) zabieranie drobnych przedmiotów (wartości mniejszej niż dwa dolary);
- 11) zabieranie rzeczy średniej wartości (między 2 a 50 dolarami);
- 12) zabieranie rzeczy dużej wartości (powyżej 50 dolarów);
- 13) zabieranie cudzych rzeczy bez wyraźnej potrzeby;
- 14) udział w bójce;
- 15) zabieranie samochodu bez zgody właściciela w celu przejechania się;
- 16) pobicie kolegi, który nic nie zawinił;
- 17) kupowanie lub spożywanie piwa, wina lub wódki;
- 18) zranienie lub zadanie bólu komuś po to, by widzieć, jak krzyczy;
- 19) celowe niszczenie lub uszkodzanie społecznego bądź prywatnego mienia;
- 20) zażywanie lub sprzedawanie narkotyków;
- 21) stosunek seksualny z osobą tej samej płci;
- 22) stosunek seksualny z osobą płci przeciwnej;
- 23) polowanie lub wędkowanie bez zezwolenia (lub naruszanie innych przepisów regulujących zasady polowania lub łowienia ryb).

Typowy zestaw odpowiedzi wyglądał następująco: „nie”, „raz lub dwa razy”, „kilkanaście razy”, „bardzo często”. Niekiedy pytano jednak bardziej dokładnie o częstość danego zachowania. Przykładowo zestaw odpowiedzi na pytanie o zawieszenie lub usunięcie ze szkoły wyglądał następująco: „nie”, „raz lub dwa razy”, „trzy lub cztery razy”, „pięć lub sześć

razy”, „ponad sześć razy”. Pytanie o ucieczki z domu miało jeszcze bardziej dokładne typy odpowiedzi: „nie”, „raz”, „dwa razy” etc. Na dalszych etapach analizy wszystkie pytania otrzymywały porównywalne, mniej dokładne typy odpowiedzi w rodzaju: „kilkanaście razy”, „bardzo często”. Odpowiedzi zamieszczane były przemiennie. Jeżeli w danym pytaniu pierwszą odpowiedzią było „nie”, to w pytaniu następnym z reguły pierwszym typem odpowiedzi było „bardzo często”. Short i Nye nie wprowadzali praktycznie ograniczeń, jeśli chodzi o okres poprzedzający badanie, co do którego miałyby odnosić się ujawnione w badaniu czyny¹⁰.

Kwestionariusz był ściśle anonimowy, a w czasie realizacji badania autorzy dokładali starań, aby zmaksymalizować poczucie bezpieczeństwa respondentów i zmniejszyć ich obawy co do ewentualnych sposobów wykorzystania wyników. Toteż niezależnie od specjalnego wstępu poprzedzającego pytania, w którym akcentowano, że wszyscy rzadziej lub częściej naruszają jakieś przepisy, starano się również stworzyć warunki sprzyjające szczerym odpowiedziom. Wśród nich wskazać można:

1) przeprowadzenie badania jednocześnie we wszystkich wybranych klasach bez wstępnego uprzedzenia;

2) usytuowanie w klasie specjalnej urny, do której badani wrzucali wypełnione kwestionariusze;

3) wyjaśnienie ustne, że cel badania jest czysto naukowy i akcentowanie całkowicie anonimowego i dobrowolnego jego charakteru.

Zastosowano kilka metod ustalania prawdomówności respondentów. Przede wszystkim w samym kwestionariuszu znajdowały się dwa pytania-pułapki (pytanie nr 3 i 8). Z dalszej analizy eliminowano tych badanych, którzy udzielili na nie niezgodnych odpowiedzi. Po drugie, nie uwzględniono kilku kwestionariuszy, w których badani przyznali się do popełnienia wszystkich czynów z maksymalną częstotliwością. Po trzecie, wyeliminowano z dalszych analiz tych respondentów, którzy na kilkunasto pozycyjnej skali „przekomformizowania” udzielili trzech lub więcej negatywnych odpowiedzi. Wreszcie nie uwzględniono kwestionariuszy tych respondentów, którzy mieli ewidentne trudności z czytaniem, w związku z czym nie zrozumieli stojącego przed nimi zadania. W sumie nie uwzględniono 1% kwestionariuszy „szkolnych”. Odsetek nie uwzględnionych kwestionariuszy w grupach z zakładu poprawczego był wyższy: odrzucono 10% kwestionariuszy chłopców i 4% kwestionariuszy dziewcząt.

¹⁰ Mówimy „w praktyce”, gdyż autorom chodziło o czyny dokonywane od czasu rozpoczęcia nauki szkolnej. Tak więc przyjęto tak długi okres poprzedzający moment badania, że w istocie można mówić o braku cezury czasowej. Jednak zarówno powyższa, jak i inne kwestie metodologiczno-techniczne nie rysują się niestety zbyt klarownie w tekstach Shorta i Nye’a. Uwaga powyższa dotyczy w szczególności liczby uwzględnianych zachowań (w zależności od badanej zbiorowości i grupy wiekowej stosowano nieco inne skale i odmienne liczby pytań) oraz zagadnienia ustalania rzetelności i trafności stworzonego przez autorów narzędzia pomiaru.

Z lektury tekstów Shorta i Nye'a nie wynika, by przeprowadzili retest. Rzetelność stosowanego narzędzia była jednak ustalana na podstawie porównywania rozkładów rezultatów, jakie uzyskano w dwóch próbach „szkolnych”. Natomiast określeniu trafności służyło skalowanie Guttmana¹¹. Ponadto autorzy porównywali rezultaty uzyskiwane przez młodzież ze szkół i zakładów poprawczych. Okazało się, że jakkolwiek zgodnie z oczekiwaniami wyniki młodzieży „przestępczej” były wyższe, to jednak profile rozkładu odpowiedzi były zbliżone. Świadczy to także o tym, że skonstruowane narzędzie mierzyło to, co miało mierzyć, to jest zachowanie dewiacyjne.

Nie sposób referować wszystkich istotniejszych ustaleń badań Shorta i Nye'a. Ograniczmy się więc tylko do przytoczenia niektórych rozkładów odpowiedzi na pytania o zachowania dewiacyjne w grupie uczniów z małych miasteczek i chłopców z zakładu poprawczego. 86,7% uczniów i 91,1% wychowanków zakładu poprawczego przyznało się do udziału w bójce. Aktów wandalizmu dokonywało 60,7% uczniów i 84,3% chłopców z zakładu poprawczego. Do utrzymywania stosunków homoseksualnych przyznało się 12,0% uczniów i 10,9% chłopców z zakładu. Największe różnice pomiędzy chłopcami ze szkół i zakładów zarysowały się, jeśli chodzi o kradzieże wszystkich uwzględnionych typów. Przykładowo: podczas gdy do kradzieży rzeczy o wartości powyżej 50 dolarów przyznało się 3,5% chłopców ze szkół, to odsetek ten dla chłopców z zakładów wynosił 90,8. Ciekawe, że różnice w odsetkach przynajęcych się do poszczególnych rodzajów zachowań dewiacyjnych w grupach dziewcząt ze szkół i z zakładów poprawczych były, generalnie rzecz biorąc, mniejsze niż w porównywanych grupach chłopców.

Zasadniczym celem badań Shorta i Nye'a było stworzenie narzędzia badawczego do dalszych badań rozmiarów zachowań dewiacyjnych w populacjach tzw. „porządnej młodzieży” (nie karanej, nie pozostającej pod dozorem kuratora). W wyniku zastosowania skalowania Guttmana stworzyli oni takie skale dla chłopców (siedmio-, i dziewięcio- lub jedenastopozycyjne, w zależności od wieku i badanej populacji — szkolnych czy „przestępczych”) oraz quasi-skale dla dziewcząt¹². Skala opracowana przez Shorta i Nye'a do końca lat sześćdziesiątych była bodaj najpowszechniej stosowanym narzędziem pomiaru nie ujawnionych zachowań dewiacyjnych młodzieży, a i obecnie ma wielu zwolenników.

¹¹ Jeżeli na podstawie pytań kwestionariusza można stworzyć ten typ skali, oznacza to, zdaniem wielu statystyków, że mamy do czynienia z jednym wymiarem badanej cechy — w tym przypadku jest nią zachowanie przestępcze.

¹² przykładowo: jedenastopunktowa skala dla chłopców obejmowała pozycje: 1, 17, 2, 10, 19, 22, 11, 6, 12, 8, 20.

Badania typu *self-report* mają i w naszym kraju swą krótką historię. Pierwszą próbę oszacowania rozmiarów kradzieży wśród młodszej młodzieży podjęli już na początku lat sześćdziesiątych Malewska i Muszyński¹³. Badanie zostało zrealizowane na ogólnopolskiej próbie losowej uczniów szóstych klas szkoły podstawowej (n=2222) za pomocą anonimowej ankiety audytoryjnej. Jego zasadniczym celem było określenie stosunku szóstoklasistów do prawa własności oraz sytuacji, w których dopuszczalne jest jego naruszenie. Malewską i Muszyńskiego interesował ponadto sposób pojmowania pewnych sytuacji jako faktów kradzieży. Zatem pytanie o to, czy badanym zdarzało się zabierać cudzą własność, było jedynie skromnym fragmentem kwestionariusza służącego zasadniczo innym celom. Niemniej było to pierwsze polskie badanie, w którym podjęto próbę zastosowania metody *self-report*. Warto więc przytoczyć rozkłady odpowiedzi na pytania: „czy zdarza się, że zabierasz cudzą własność?” oraz „jak często zdarza ci się zabierać cudzą własność?”¹⁴. 0,8% dzieci odpowiedziało, że zdarza się im to „bardzo często”, 4,2% stwierdziło, że zdarza się to im „często”, 26,6% badanych dzieci przyznało się, że zdarza się im kraść „niekiedy”, 34,5% zabierało cudzą własność „rzadko”. W odpowiedzi na pytanie o częstość kradzieży 28,7% badanych dzieci stwierdziło, że nigdy nie zabierały cudzej własności¹⁵. Na uwagę zasługuje fakt, że odsetek negatywnych odpowiedzi na pytanie sformułowane bardziej ogólnie, tj. „czy zdarza się, że...” był wyższy o 13. Świadczy to, że w kwestionariuszu *self-report* korzystnie; jest formułować bardziej konkretne pytania o zachowania dewiacyjne. Zagadnienie to zostanie szerzej zanalizowane w innym miejscu.

Jako przykład badania typu *self-report* przytaczana bywa praca Szemińskiej i Gołąba¹⁶. Jednak lektura pracy nasuwa przypuszczenie, że niewiele ma ona wspólnego z nurtem badań typu *self-report*, zresztą nie to stanowi główny przedmiot ich zainteresowań. Bowiern Szemińska i Gołąb, podobnie jak Malewska i Muszyński, dążyli do określenia „poczucia moralnego” grup młodzieży ze szkół i zakładu wychowawczego, przede wszystkim stosunku do normy „nie kradnij”. Dlatego też kwestia rozmiarów zachowań dewiacyjnych czy przestępczych potraktowana została bardziej niż marginalnie. Autorzy przebadali 64 nieletnich przestępców z trzech warszawskich zakładów wychowawczych oraz 61 uczniów siódmej klasy szkoły

¹³ H. Malewska: *Norma uczciwości w środowisku młodzieży. Refleksje nad przestrzeganiem tej normy oparte na niektórych wynikach badań ankietowych*, „Studia Socjologiczne”, 1963, nr 2; H. Muszyński: *Kradzież w poglądach i postępkach dzieci*, „Nowa Szkoła”, 1963, nr 1; H. Malewska, H. Muszyński: *Children's Attitudes to Theft*. „The Polish Sociological Bulletin”, 1964, no. 1.

¹⁴ H. Muszyński: *op. cit.*

¹⁵ H. Muszyński: *op. cit.*

¹⁶ A. Szemińska, A. Gołąb: *Norma „nie kradnij” w świadomości moralnej młodzieży przestępczej i nieprzestępczej*, „Psychologia Wychowawcza”, 1969, nr 4.

podstawowej¹⁷. Szemińska i Gołąb wspominają, że „badani wypowiadali się na piśmie w warunkach anonimowych”. Lakoniczne uwagi autorów na temat rozmiarów zachowań dewiacyjnych badanych można ująć następująco: Chłopcy z porównywanych grup różnili się istotnie pod względem częstości dokonywania kradzieży, jak też wagi tych czynów. Podczas gdy większość przestępców przyznała się do popełnienia wielu kradzieży¹⁸, to nieprzestępcy „w ogromnej większości przyznali się do najwyżej 1—3 kradzieży najczęściej drobnych przedmiotów zabieranych kolegom szkolnym lub członkom najbliższej rodziny z myślą o osobistym użytkowaniu skradzionych rzeczy”¹⁹. Jest dyskusyjne, czy fakt kradzieży motywowany chęcią „osobistego użytkowania” skradzionych przedmiotów jest wystarczającą przesłanką upoważniającą do określania kradzieży dokonywanych przez młodzież szkolną jako mniej poważnych, jak czynią to autorzy, przeciwstawiając ten typ motywacji „chęci zdobycia pieniędzy”, która jakoby stanowiła dominujący motyw kradzieży dokonywanej przez chłopców z zakładów wychowawczych.

Szemińska i Gołąb najwyraźniej nie uświadamiali sobie konsekwencji, jakie wynikają ze stwierdzenia faktu, że zarówno młodzież „nieprzestępcza”, jak i „przestępcza” dokonuje kradzieży. W istocie bowiem obydwie grupy były „przestępcze”. W dalszych rozważaniach jednak Szemińska i Gołąb starają się dowieść, że między porównywanymi grupami istnieją zasadnicze różnice pod względem „poczucia moralnego”. Przy czym, wbrew wcześniej poczynionym przez siebie ustaleniom, operują przeciwstawną parą pojęć — przestępcy i nieprzestępcy — bez stosowania koniecznego w tej sytuacji cudzysłowu.

Generalnie rzecz biorąc metodologiczna warstwa badania Szemińskiej i Gołąba pozostawia wiele do życzenia²⁰, co sprawia, że ich ustalenia, i bez tego niejednoznaczne, mogą być zasadnie kwestionowane. Szkoda, że autorzy nie przytoczyli dokładniejszych danych ilustrujących rozmiary uwzględnionych zachowań dewiacyjnych w analizowanych zbiorowościach.

W latach 1976—1977 Ostrianska i Wójcik²¹ przeprowadziły duże badanie uczniów III—VIII klas szkół podstawowych z terenu Warszawy. Do ba-

¹⁷ Autorzy nie podają niestety bliższych informacji o kryteriach i sposobie doboru próby oraz technice prowadzenia badania.

¹⁸ A. Szemińska i A. Gołąb nie przytaczają dokładnych danych liczbowych, stwierdzają tylko, że liczba ta wynosiła od kilkunastu do ponad stu.

¹⁹ A. Szemińska i A. Gołąb: *op. cit.* s. 462.

²⁰ W szczególności dotyczy to sposobu grupowania danych i posługiwania się technikami statystycznymi. Np. dokonuje się tu określenia poziomu istotności zaobserwowanych różnic mimo nielosowego charakteru próby.

²¹ Z. Ostrianska, D. Wójcik: *Raport z badań nad rozmiarami i uwarunkowaniami nieprzystosowania społecznego u uczniów warszawskich szkół podstawowych* (tekst powielony); Z. Ostrianska: *Norma poszanowania cudzej własności w zachowaniu uczniów szkół podstawowych*, „Państwo i Prawo”, 1980, nr 7.

dania wylosowano 50 szkół, następnie zaś w szkołach tych dokonano losowania 120 klas, w których przeprowadzono badanie przy zastosowaniu anonimowej ankiety audytoryjnej (n=3177, w tym 1631 chłopców i 1546 dziewcząt). Badanie typu *self-report* stanowiło część szerszego programu badawczego, którego celem było oszacowanie rozmiarów nieprzystosowania społecznego młodzieży w młodszym wieku i ustalenie jego uwarunkowań.

W ankiecie pytano o: 1) niepowodzenia w nauce szkolnej, 2) wagary, 3) ucieczki z domu, 4) picie alkoholu, 5) zażywanie narkotyków, 6) jazdę tramwajem lub autobusem bez biletu, 7) niszczenie cudzego mienia, 8) inne czyny przeciwko mieniu (w tym kradzieże, oszustwa, rozbój — „zabranie czegoś młodszemu dziecku siłą”). Niezależnie od powyższych grup pytań o zachowania dewiacyjne autorki zamieściły w swym kwestionariuszu pytania o to, czy badani byli zatrzymani przez MO w związku z podejrzeniem o kradzież („zabranie cudzej własności”) i czy mieli sprawę w sądzie dla nieletnich.

Zestaw odpowiedzi przedstawiał się następująco: „nigdy”, „1 raz”, „2—3 razy”, „4—10 razy”, „więcej razy”.

Oto najistotniejsze ustalenia badania Ostrihanskiej i Wójcik. Zgodnie z oczekiwaniami tak poziom, jak i nasilenie (częstotliwość dokonywania) zachowań dewiacyjnych były w grupie chłopców wyższe niż w grupie dziewcząt. Przykładowo: w grupie najstarszych chłopców (15 lat) jedynie 15,3% nie przyznało się do faktu zabierania cudzej własności, to w grupie najstarszych dziewcząt odsetek nie dokonujących kradzieży był wyższy przeszło o połowę (38,0%). W tej grupie wiekowej 16,3% chłopców dokonywało kilkunastokrotnych lub jeszcze częstszych kradzieży („więcej niż 10 razy”), podczas gdy odsetek dziewcząt dokonujących wielokrotnych kradzieży wynosił zaledwie 1,4. Do niejednokrotnych kradzieży przyznało się ogółem 48% chłopców, podczas gdy odsetek dziewcząt wśród „recydywistów” wynosił tylko 26. Natomiast, co ciekawe, odsetki dziewcząt i chłopców przyznających się do jednokrotnego zabierania cudzej własności były bardzo podobne (z wyjątkiem najstarszych grup wiekowych). Jeśli idzie o typy kradzieży wyróżnione ze względu na miejsce dokonania czy osobę ofiary, to w grupie chłopców (w referowanym tekście brak jest analogicznych danych dotyczących dziewcząt) najczęstsze były kradzieże z cudzych działek i ogrodów — 35,2%, następnie okradanie rodziców — 22,3% oraz kradzieże ze sklepów samoobsługowych — 18,1%. W badanej zbiorowości najrzadziej pojawiały się kradzieże z piwnic (5,8%) i kradzieże wina ze sklepu (0,7%). Natomiast już kradzieże słodyczy ze sklepu zdarzały się częściej (11,2%).

W badaniu Ostrihanskiej i Wójcik na uwagę zasługuje fakt, że podobnie jak uprzednio w badaniu Malewskiej i Muszyńskiego, na pytanie ogólne o to, czy badanym „zdarzało się kiedyś zabierać cudzą własność”, jedynie

13,7% chłopców udzieliło pozytywnej odpowiedzi, podczas gdy rozkład odpowiedzi na pytania szczegółowe wskazywał, że w rzeczywistości kradnący chłopcy stanowili blisko 70%.

Zabieranie cudzej własności plasowało się na czwartym miejscu na liście uwzględnionych przez autorki zachowań dewiacyjnych, jeśli idzie o częstotliwość dokonywania (w grupie chłopców). Na pierwszym miejscu znalazło się okłamywanie (88,8% odpowiedzi pozytywnych w grupie chłopców), co w dużym stopniu potwierdza wiarygodność odpowiedzi na pytanie ankiety. Bowiem jest to istotnie nader częste zjawisko. Blisko 90% badanych chłopców przyznało się także do odpisywania lekcji od kolegów, co świadczy, że większość badanych przyznaje się najczęściej do najbardziej błahych i rozpowszechnionych przewinień. Z innych nagannych lub (i) karalnych zachowań 25,2% badanych chłopców dokonywało aktów wandalizmu, 2,8% uciekało z domu, 2,4% zaś zażywało narkotyki. Wystąpiła tendencja wzrostowa z wiekiem: starsi chłopcy przyznawali się do większej liczby zachowań dewiacyjnych. Przykładowo, w grupie piętnastolatków 84% przyznało

się do niejednokrotnych kradzieży, w tym 1/5 stanowili sprawcy kradzieży częstszych niż dziesięciokrotne.

Jest rzeczą charakterystyczną, że badania Ostrianskiej i Wójcik ujawniły dużą dysproporcję pomiędzy odsetkami chłopców, którzy przyznali się do dokonania kradzieży (w tym także i „wielokrotnych recydywistów”), a odsetkami badanych, którzy przyznali się do tego, że byli zatrzymani przez MO w związku z podejrzeniem o dokonanie kradzieży. Przykładowo: do zatrzymania w związku z podejrzeniem o kradzież przyznało się 0,3% najmłodszych chłopców (9—10 lat) i 10% chłopców najstarszych (15 lat). Jeszcze niższy odsetek badanych miał sprawę w sądzie: odpowiednio 0,1% i 7,4%. Dane te w sposób nader wymowny ilustrują tezę o niewielkiej adekwatności urzędowych danych dotyczących rozmiarów zachowań dewiacyjnych w stosunku do ich rozmiarów faktycznych. Świadczą ponadto o tym, jak selektywnie działa aparat ścigania i wymiaru sprawiedliwości i jak zawodne mogą być rezultaty badań prowadzonych na osobach „urzędowo” uznanych za „nieletnich przestępców”. Zdaniem Ostrianskiej sprawcy niejednokrotnych kradzieży stanowili, począwszy od trzynastego roku życia,

większość wśród badanych uczniów, w tym przeszło 1/3 to sprawcy co najmniej czterokrotnych kradzieży.

Najwszechstronniejsze badanie typu *self-report* i najbardziej zbliżone do innych tego typu badań prowadzonych na świecie przeprowadzili Ostrowska i Siemaszko²². Objęto nim 2994 osoby, w tym 1196 dziewcząt i 1795 chłopców — uczniów warszawskich szkół ponadpodstawowych wszy-

²² K. Ostrowska, A. Siemaszko: *Rozmiary spożywania napojów alkoholowych przez młodzież szkół ponadpodstawowych na terenie Warszawy w 1979 roku*, „Studia Kryminologiczne, Kryminalistyczne i Penitencjarne”, 1983, t. XIII;

stkich głównych rodzajów (nie uwzględniono szkół artystycznych i policealnych). Brak zgody władz oświatowych uniemożliwił dobór próby metodą losowania. Dlatego też zastosowano dobór celowy. Aby zmniejszyć niedostatki wynikające z takiego sposobu doboru próby, starano się wybierać szkoły z różnych obszarów Warszawy (zarówno z centrum, jak i z peryferii) i o zróżnicowanym poziomie nauczania (szkoły „elitarne”, o wysokim poziomie nauczania i słabsze pod tym względem). O ile było to możliwe, starano się również eliminować szkoły, w których już uprzednio prowadzono badania ankietowe. Wynikało to z obawy, czy uprzedni „trening” nie wpłynie na rzetelność odpowiedzi.

Odsetki młodzieży z poszczególnych typów szkół przedstawiały się następująco: 19,7% — zasadnicze szkoły zawodowe, 28,7% — licea ogólnokształcące, 28,7% — licea zawodowe, 25,3% — technika.

Badaniem objęto młodzież ze wszystkich klas pierwszych i ostatnich w wybranych szkołach. Chodziło bowiem o wychwycenie ewentualnych różnic w rozmiarach zachowań dewiacyjnych między młodzieżą w młodszym i starszym wieku. Badanie zostało przeprowadzone za pomocą anonimowej, dobrowolnej ankiety audytoryjnej. Starano się stworzyć warunki sprzyjające szczerości wypowiedzi. Między innymi przyjęto zasadę, że nauczyciele nie mogą przebywać w klasie w trakcie badania. Realizacja badania przebiegała równolegle we wszystkich klasach pierwszych bądź ostatnich. Miało to na celu uniemożliwienie respondentom porozumienia się, np. w trakcie przerw, i ewentualną wymianę uwag co do charakteru i celu badania.

W trzy miesiące po badaniu właściwym dokonano retestu, którym objęto młodzież ze wszystkich klas „A”. Rezultaty retestu świadczyły o nadspodziewanie dużej stałości wyników w czasie. Żadna z różnic rozkładów odpowiedzi na pytania o zachowania dewiacyjne między badaniem właściwym a retestem nie była statystycznie istotna.

Kwestionariusz zawierał 42 pytania o różnorodne zachowania dewiacyjne, przekroczenia przepisów porządkowych, wykroczenia i przestępstwa. Od bardzo błahych (w rodzaju „jazdy na gapę” czy „niezwracanie reszty z zakupów”) do bardzo poważnych, jak paserstwo, włamanie czy rozbój. Wszystkie pytania miały jednakowy zestaw odpowiedzi: „nigdy”, „raz czy dwa razy”, „kilka razy”, „kilkanaście razy”, „częściej”.

Ogólnie rzecz biorąc badaną młodzież cechował nadspodziewanie duży poziom zachowań dewiacyjnych. Przykładowo: w grupie chłopców 539 osób, tj. 30,9% przyznało się do popełnienia przynajmniej jednokrotnie połowy z 42 czynów uwzględnionych w kwestionariuszu, z czego 2,8% dokonało od 27 do 31 czynów, 2,3% zaś od 32 do 42 czynów. Zatem łącznie 6,1%

A. Siemaszko: *Rozmiary i uwarunkowania zachowań dewiacyjnych uczniów szkół ponadpodstawowych z terenu Warszawy*, „Studia Kryminologiczne, Kryminologiczne i Penitencjarne”, 1983, t. XIV.

badanych chłopców cechował wyjątkowo wysoki poziom dewiacji. Odrębne analizy wykazały, że istnieje ścisła zależność między liczbą popełnianych czynów a częstością ich popełniania. Zatem wspomniane 6% badanych chłopców (300 osób) można śmiało określić mianem „wielokrotnych recydywistów” w pojęciu stosowanym w badaniach typu *self-report*.

Analogicznie jak w badaniach Ostrianskiej i Wójcik także w badaniach Ostrowskiej i Siemaszki ujawniły się spore różnice w poziomie zachowań dewiacyjnych w zależności od wieku i płci. Przykładowo: poziom dewiacji w grupie starszych chłopców był przeciętnie czterokrotnie wyższy niż w grupie młodszych dziewcząt.

Wśród najbardziej rozpowszechnionych zachowań znalazły się m.in. jazda „na gapę” (94,6% odpowiedzi pozytywnych w grupie dziewcząt i 96,1% w grupie chłopców), niezwracanie reszty z zakupów (odpowiednio: 79,6% i 84,1%), niezwrócenie reszty sprzedawcy, jeżeli omyłkowo wydał jej za dużo (67,6% i 84,0%) oraz nieoddanie znalezionej rzeczy (69,8% i 83,8%). Stosunkowo częste były czyny agresywne, szczególnie w grupie chłopców. „Bez wyraźnej przyczyny” pobiło kogoś 20,2% chłopców i 6,5% dziewcząt, co i tak jest zaskakujące, jeśli zważyć rodzaj inkryminowanego zachowania. 11,6% badanych chłopców i 2,8% dziewcząt brało udział w bójce, w której stosowano niebezpieczne narzędzia. Wśród kradzieży dominowały stosunkowo mniej poważne. I tak 16,9% dziewcząt oraz 31,6% chłopców przyznało się do kradzieży rzeczy o wartości mniejszej niż sto złotych, 20,8% dziewcząt i 42,4% chłopców dokonywało kradzieży sklepowych, 8,2% dziewcząt i 14,6% chłopców okradało różnego rodzaju automaty. Poważną kradzież, określoną przez autorów jako kradzież rzeczy lub pieniędzy o wartości przewyższającej 1000 zł (badanie było realizowane w roku szkolnym 1979/1980) popełniło 1,6% dziewcząt i 4,6% chłopców, co stanowiło najniższe odsetki spośród wszystkich uwzględnionych przez autorów zachowań dewiacyjnych.

Na zakończenie skrótowej prezentacji wyników badania Ostrowskiej i Siemaszki przedstawiamy rozkład odpowiedzi na pytania o zachowania, które — choć nie stanowią przestępstw ani wykroczeń — są powszechnie uznawane za przejawy nieprzystosowania społecznego lub zwiastuny demoralizacji. Do zachowań takich należą bez wątpienia ucieczki z domu, do których przyznało się 5% dziewcząt i 8,5% chłopców; 3,8% dziewcząt i 13,7% chłopców było zawieszonych w prawach ucznia. Do zażywania narkotyków i środków odurzających przyznało się 10,5% dziewcząt i 6,3% chłopców, były to jednak w przytłaczającej większości jednorazowe eksperymenty.

Referowane uprzednio badanie zostało powtórzone w 1981 r.²³ na losowej próbie młodzieży ze szkół ponadpodstawowych na terenie pięciu

²³A. Siemaszko: *Rozmiary zachowań dewiacyjnych młodzieży wiejskiej*, „Studia Kryminologiczne, Kryminalistyczne i Penitencjarne”, 1986, t. XVIII; e a de m: *Rozmiary i uwarunkowania zachowań dewiacyjnych młodzieży wiejskiej*, „Archiwum Kryminologii”, 1986, t. XIII.

województw o charakterze typowo rolniczym²⁴. Badanie zrealizowano w szkołach mieszczących się na terenie województw: bielskopodlaskiego, łomżyńskiego, ostrołęckiego, siedleckiego i zamojskiego. Objęto nim 2144 osoby, w tym 1702 chłopców i 429 dziewcząt w wieku od 14 do 19 lat z 29 szkół²⁵. W badaniu wzięła udział młodzież z 86 klas. W części dotyczącej zachowań dewiacyjnych zastosowano identyczny kwestionariusz jak w poprzednim badaniu. Autorzy zmierzali bowiem do ustalenia, czy i w jakim stopniu rozmiary zachowań dewiacyjnych młodzieży, którą z pewną dozą umowności określić można jako wiejską, różnią się od struktury i rozmiarów zachowań dewiacyjnych młodzieży wielkomiejskiej. Jak wiadomo, informacje o zachowaniach dewiacyjnych młodzieży zamieszkałej na wsi lub w małych miasteczkach są bardziej niż skąpe. Referowane badanie zmierzało zatem m.in. do wypełniania tej luki.

Analogicznie jak w badaniu prowadzonym uprzednio w Warszawie Ostrowska i Siemaszko zastosowali anonimową ankietę audytoryjną. Mimo dużych różnic obydwu badanych zbiorowości rezultaty badania „wieskiego”, okazały się zaskakująco zbieżne z wynikami badania „warszawskiego”, co w dużym stopniu świadczy o rzetelności skonstruowanego narzędzia pomiaru zachowań dewiacyjnych.

Najistotniejsze ustalenia badania Ostrowskiej i Siemaszki przedstawiają się następująco. W grupie zachowań określonych umownie jako niesubordynacja (m.in. niewracanie na noc bez zgody rodziców, palenie papierosów przed 14 rokiem życia, zawieszenie w prawach ucznia itp.) najpowszechniejsze było palenie papierosów. Do palenia papierosów przed 14 rokiem życia przyznało się 78,2% chłopców i 44,8% dziewcząt. Do ucieczki z domu przyznało się niecałe 10% respondentów, z czego 2% uciekało z domu kilkakrotnie. Dziewczęta przyznawały się dwukrotnie rzadziej do ucieczek z domu (5%) niż chłopcy (10%). Przeszło 20% respondentów przyznało się do faktu legitymowania przez milicję.

Zaznaczały się statystycznie istotne różnice między chłopcami i dziewczętami. Podczas gdy do legitymowania przyznało się niecałe 10% dziewcząt, to odsetek legitymowanych chłopców wyniósł blisko 30, w grupie najstarszych chłopców był jeszcze wyższy (45%). Rozkład odpowiedzi na pytanie o zażywanie narkotyków przedstawiał się następująco: „nigdy” — 93,8% (dziewczęta — 95,1%, chłopcy — 93,5%), „raz czy dwa razy” — 4,3% (dziewczęta 3,8%, chłopcy — 4,4%), „kilka razy — 1,1% (dziewczęta — 0,5%, chłopcy — 1,3%), „kilkanaście razy” — 0,9% (dziewczęta — 0,7%,

²⁴ Przy określaniu rolniczego charakteru województwa przyjęto następujące kryteria. 1) niski odsetek ludności utrzymującej się głównie ze źródeł pozarolniczych, 2) niski odsetek ludności zamieszkałej w miastach, 3) mała wartość środków produkcji, 4) niewielka liczba miast powyżej 10 tys. mieszkańców, 5) niski odsetek ludności zatrudnionej w gospodarce społecznoniej.

²⁵ Ponieważ próba była losowana ze względu na liczbę chłopców, stąd zarysowała się wyraźna dysproporcja płci.

chłopcy ___0,8%). Wiek respondentów nie wprowadzał do powyższego rozkładu korekt tak w grupie chłopców, jak i dziewcząt.

W grupie zachowań „nieuczciwych” najbardziej rozpowszechniona była jazda „na gapę”. Jedynie 17% badanych nie przyznało się do jazdy bez ważnego biletu. Dziewczęta były pod tym względem znacznie bardziej uczciwe od chłopców. Ponadto zarysowała się wyraźna tendencja do częstszego przyznawania się do jazdy bez ważnego biletu wraz z wiekiem. Przykładowo: w grupie najstarszych chłopców już blisko 90% przyznało się do jazdy „na gapę”, z czego 15% czyniło to notorycznie — kilkanaście razy lub częściej.

Blisko połowa badanych przyznała się do niezwrócenia pożyczonych rzeczy, przy czym 40% respondentów postąpiło tak „raz czy dwa razy” , 7% zaś częściej. Analogicznie jak w przypadku innych zachowań nagananych bądź przestępczych, także i w tym przypadku zarysowała się tendencja wzrostowa z wiekiem. I tak np. w grupie najmłodszych (14 —15 lat) chłopców do nieoddawania pożyczonych rzeczy przyznało się 37,2% badanych, natomiast w grupie najstarszych chłopców (powyżej 17 lat) odsetek ten wynosił już 56,9.

Do tej grupy zachowań zaliczono również zabieranie rodzicom pieniędzy bez ich wiedzy, choć równie dobrze czyn ten mógłby się znaleźć w kategorii przestępstw i wykroczeń. Blisko połowa badanych przyznała się do kradzieży pieniędzy rodzicom, z czego 26,1% ukradło „raz czy dwa razy” , 15,1% — „kilka razy”, a 5,6% — „kilkanaście razy” lub jeszcze częściej. Wystąpiły duże różnice w przekroju płci i wieku. Przykładowo, w grupie najstarszych chłopców odsetek zabierających pieniądze rodzicom bez ich wiedzy był blisko 20 wyższy niż w grupie najmłodszych dziewcząt i wynosił 52,9.

W kategorii zachowań nazwanych przez autorów z pewną umownością „przestępstwami” wyróżniono czyny przeciwko osobie, mieniu oraz pozostałe.

Wśród czynów o charakterze agresywnym dominowały bójki (45% odpowiedzi pozytywnych w całej grupie) i pobicia (50% odpowiedzi pozytywnych w całej grupie). Z 45% respondentów, którzy brali udział w bójce, blisko 30% odpowiedziało, że zdarzyło im się to „raz czy dwa razy”, niecałe 15% badanych brało udział w bójce kilkakrotnie. Podczas gdy blisko

3/4dziewcząt zadeklarowało, że nigdy nie brało udziału w bójce, to u chłopców odsetek uczestniczących w bójkach sięgał 50 (w grupie najstarszych chłopców wynosił 54,2).

Na pytanie o pobicie chuligańskie, które zostało zoperacjonalizowane jako „pobicie kogoś dla zwykłej fantazji”, odpowiedzi pozytywnych w całej grupie było blisko 1/4, przy czym około 7% dopuściło się pobicia o charak-

terze chuligańskim co najmniej kilkakrotnie. Zgodnie z oczekiwaniami odsetek dziewcząt przyznających się do tego czynu był znacznie mniejszy niż chłopców (odpowiednio 9,4 i 28,1), choć i tak można go uznać za nadspodziewanie wysoki. W przeciwieństwie do innych rezultatów wiek nie zróżnicował w tym przypadku grupy chłopców.

Wśród przestępstw przeciwko mieniu dominowały drobne kradzieże. Do kradzieży rzeczy lub pieniędzy o wartości mniejszej niż 100 zł przyznało się 24,9% dziewcząt i 32,4% chłopców, w grupie najstarszych chłopców zaś odsetek ten wynosił 38,6. Zarówno w grupie dziewcząt, jak i chłopców blisko 20% dokonywało takich kradzieży co najmniej kilkakrotnie. Blisko 25% badanych przyznało się do dokonania kradzieży ze sklepu. Również i w tym przypadku ujawniły się spore różnice na niekorzyść chłopców. Odsetek kradnących ze sklepu wynosił w grupie chłopców blisko 30 (w grupie najstarszych chłopców 41,6), w grupie zaś dziewcząt nieco ponad 14. Włamań do komórek, piwnic, strychów etc. dokonywało 15% badanych, w tym 6,1% dziewcząt i 17,2% chłopców, z których niecałe 5% stanowili „wielokrotni recydywiści”.

Wśród badanej młodzieży nadspodziewanie częste było paserstwo. Na pytanie, „czy zdarzyło ci się kupować lub dostawać rzeczy, o których wiedziałeś, że pochodzą z kradzieży?”, 1/4 respondentów udzieliła pozytywnej

odpowiedzi, w tym jedynie 7,3% dziewcząt i aż 29,9% chłopców. Odsetek w grupie najstarszych chłopców wynosił 42,2, przy czym 15% przyznało, że postępowo tak wielokrotnie.

W omawianej zbiorowości najrzadziej przyznawano się do następujących czynów o znamionach przestępstw: rozbój (2,4% odpowiedzi pozytywnych w grupie dziewcząt i 10,1% w grupie chłopców), okradanie automatów (6,6% odpowiedzi pozytywnych w grupie dziewcząt i 7,8% w grupie chłopców), kradzież rzeczy lub pieniędzy o wartości 500—1000 zł (6,2% odpowiedzi pozytywnych w grupie dziewcząt i 7% w grupie chłopców), wyjście z lokalu bez zapłacenia rachunku (3,3% odpowiedzi pozytywnych w grupie dziewcząt i 5,8% w grupie chłopców).

Najradszym przestępstwem była kradzież rzeczy lub pieniędzy o wartości przekraczającej 1000 zł. Odsetek odpowiedzi pozytywnych wynosił w całej grupie około 5, z czego 1% badanych dokonywało takiej kradzieży kilkakrotnie lub jeszcze częściej. Podczas gdy w grupie dziewcząt do kradzieży rzeczy lub pieniędzy o wartości ponad 1000 zł przyznało się 2,8%, to w grupie chłopców analogiczny odsetek wynosił już 5,6, w grupie zaś najstarszych chłopców był jeszcze wyższy — 8,3. Wśród najstarszych chłopców tę stosunkowo poważną kradzież 6% popełniło „raz czy dwa razy”, natomiast 2,2% kilkakrotnie lub jeszcze częściej.

Rezultaty referowanych powyżej polskich badań, w których wykorzystano metodę *self-report* do szacowania rozmiarów zachowań dewiacyj-

nych uczącej się młodzieży, nasuwają następujące wnioski. W płaszczyźnie teoretycznej ich wyniki upoważniają do konstatacji, iż dewiacja czy przestępczość jest cechą ilościową, a nie jakościową. Badana „porządna”, ucząca się młodzież wykazywała znaczne rozmiary zachowań dewiacyjnych, w tym *stricte* przestępczych. Od młodzieży uznanej urzędowo za „nieletnich przestępców” młodzież ta różni się strukturą i liczbą popełnianych przestępstw, nie są to jednak zbiorowości odmienne „jakościowo”. Ma to, jak się wydaje, kapitalne znaczenie w sferze konstruowania i weryfikowania koncepcji teoretycznych w kryminologii i socjologii dewiacji, gdzie zazwyczaj operuje się zdychotomizowaną parą pojęć: przestępcy i nieprzestępcy (czy też dewianci i niedewianci). W świetle rezultatów referowanych badań podejście takie nie wyda je się uprawnione.

W płaszczyźnie praktycznej rezultaty analizowanych badań świadczą, iż za pomocą kwestionariuszy typu *self-report* można uzyskiwać wiarygodne informacje o zachowaniach dewiacyjnych. Zatem ten typ narzędzia badawczego może służyć zarówno celom ściśle naukowym, jak również można by je wykorzystywać w szacowaniu rozmiarów „ciemnej liczby przestępstw. Okazuje się ponadto, że uczącą się, teoretycznie „porządną” młodzież cechują wcale niemałe rozmiary dewiacji, większe niż można było przypuszczać. Na przykład badania Ostrowskiej i Siemaszki wskazują, że blisko 10% respondentów z obydwu prób wykazywało wysoki stopień dewiacyjnego zaangażowania.

Wreszcie w płaszczyźnie metodologicznej wyniki polskich badań typu *self-report* stawiają pod znakiem zapytania rozpowszechniony schemat badawczy, w którym jako grupę kontrolną w stosunku do nieletnich czy młodocianych przestępców przyjmuje się młodzież szkolną. Zakłada się przy tym, że postępuje ona zgodnie z prawem. Tymczasem prawa przestrzega co najwyżej część tej młodzieży. Należałoby więc w ogóle zrezygnować z takiego schematu badawczego albo też wprowadzić doń istotną modyfikację. Polegałaby ona na tym, by zamiast dwóch grup — badawczej i kontrolnej — schematy badań opierać na doborze trzech takich grup. Pierwszą stanowiłaby młodzież uznana za przestępczą przez formalne instytucje kontroli społecznej (prokuratura, milicja, sąd dla nieletnich). Drugą — młodzież „nieoficjalnie” przestępcza, której przestępczości nie ujawniono, choć stwierdzono za pomocą badań *self-report*. I wreszcie w wyniku selekcji rezultatów uzyskiwanych przez respondentów na skalach dewiacji czy przestępczości otrzymywano by prawidłową grupę kontrolną. Proponowany schemat badawczy spełniałby jednocześnie postulat traktowania dewiacji jako zjawiska ilościowego, ponieważ każda z rozpatrywanych grup prezentowałaby inny stopień dewiacyjnego zaangażowania.

Dotychczasowe badania *self-report* przeprowadzane w naszym kraju nasuwają następujące szczegółowe uwagi natury metodologicznej.

Po pierwsze, świadczą one, że technika *self-report* może być także i w

naszym kraju wykorzystywana z powodzeniem w badaniach nie ujawnionych zachowań dewiacyjnych.

Po drugie, mimo odmiennych zbiorowości, które badano, i poważnych różnic w stosowanych narzędziach badawczych (które generalnie czynią porównania wyników badań *self-report* wyjątkowo trudnym zadaniem) uzyskano uderzające zbieżności np. co do rozmiarów kradzieży.

Po trzecie, należałoby w przyszłości starać się o stworzenie uniwersalnego narzędzia badawczego, które by mogło być wykorzystywane w szacowaniu rozmiarów zachowań dewiacyjnych w różnych populacjach o odmiennych parametrach. Przykładowo: wszystkie prezentowane badania (z wyjątkiem fragmentu badania Szemińskiej i Gołęba) były prowadzone na próbach uczącej się młodzieży.

O ile w przypadku badań młodzieży ze szkół podstawowych prawdopodobieństwo, że poza próbą znajdzie się jakaś większa frakcja młodzieży, jest minimalne ze względu na obowiązek szkolny, to w badaniach starszej, uczącej się młodzieży reprezentatywność prób w stosunku do populacji młodzieży w tym wieku stanowi istotny problem. Dlatego też wskazane byłoby przeprowadzenie także badań młodzieży pracującej.

Po czwarte, jakkolwiek dotychczasowe wyniki badań *self-report* prowadzone w Polsce przemawiają na rzecz trafności i rzetelności tego narzędzia pomiaru, są to jednak w dalszym ciągu w większej mierze „ogólne wrażenia”, intuicje, poparte głównie analizą trafności treściowej. Jeden retest, i to niezbyt metodologicznie wyrafinowany, stanowi dość wątplą podstawę do określenia rzetelności skal do badania nie ujawnionych zachowań dewiacyjnych. Wydaje się zatem celowe skoncentrowanie przyszłych wysiłków na metodologicznej stronie badań *self-report*, bowiem zunifikowane, rzetelne i trafne narzędzie pomiaru zachowań dewiacyjnych młodzieży mogłoby być bardzo pomocne we wszelkich badaniach nieprzystosowania społecznego dzieci i młodzieży, zmuszając zapewne do weryfikacji wielu, zdawałoby się, niepodważalnych ustaleń co do jego etiologii. Wymagałoby to jednak rozwiązania bardzo trudnych problemów metodologicznych, z których najważniejsze zostaną poniżej scharakteryzowane.

III. WYBRANE ZAGADNIENIA METODOLOGICZNE BADAŃ TYPU *SELF-REPORT*

I. UWAGI WPROWADZAJĄCE

W rozważaniach poniższych zajmiemy się najistotniejszymi kwestiami natury metodologicznej, które pojawiają się najczęściej w badaniach typu *self-report*. Ramy opracowania uniemożliwiają, rzecz prosta, analizę wszystkich zagadnień związanych ze stosowaniem metody typu *self-report* w badaniach zachowań dewiacyjnych. Dlatego też skoncentrowaliśmy się

na problemach naszym zdaniem najistotniejszych, czyli takich, od których sposobu rozwiązania zależy w największym stopniu uzyskanie szczerych i dokładnych odpowiedzi.

Do analizy wybrano następujące zagadnienia: sposób formułowania pytań, ich liczba oraz rodzaje zestawów odpowiedzi, sposób rozwiązywania kwestii cezury czasowej między chwilą dokonania danego czynu a momentem realizacji badania, problem ważenia zachowań dewiacyjnych ze względu na ich rozmaity „ładunek dewiacji”, rodzaje badań typu *self-report* i wreszcie kluczową w badaniach typu *self-report* sprawę rzetelności i trafności stosowanego narzędzia pomiaru. Uznać ją należy za kluczową, gdyż wątpliwości co do prawdomówności respondentów stanowiły przez długi czas jedną z poważniejszych przeszkód na drodze do całkowitej akceptacji metody *self-report* jako sposobu uzyskiwania informacji o nie ujawnionych zachowaniach dewiacyjnych. W podsumowaniu rozważymy obecny stan metodologii badań typu *self-report* oraz możliwości i kierunki ewentualnego udoskonalania tego narzędzia pomiaru.

2. SPOSOBY FORMUŁOWANIA PYTAŃ

W badaniach typu *self-report* istotne znaczenie ma sposób, w jaki zadawane są pytania dotyczące zachowań dewiacyjnych. Od zakresu ogólności pytania bowiem, jego większej lub mniejszej eufemistyczności, stopnia, w jakim sprecyzowana jest sytuacja, w której działał ewentualny „sprawca” etc., może zależeć w niemałym stopniu tak właściwe zrozumienie pytania, jak też dokładność i szczerść odpowiedzi. Z wielu kwestii, jakie wiążą się ze sposobem formułowania pytań o zachowania dewiacyjne, zasygnalizujemy poniżej najważniejsze.

Badacz, który przystępuje do konstruowania kwestionariusza *self-report*, musi przede wszystkim rozstrzygnąć problem, jak ujmować pytania: mówiąc wprost o kradzieży czy też wyrażając się eufemistycznie? Innymi słowy: czy pytanie ma mieć formę „czy zdarzało ci się ukraść?”, czy też „czy zdarzało ci się zabrać i nie oddać?”. Wbrew pozorom niełatwo jest rozstrzygnąć ten dylemat. Intuicja podpowiada, aby formułować pytania w sposób eufemistyczny, by nie wywoływać u respondentów nieprzyjemnych skojarzeń, jakie może nasuwać słowo „kradzież” w kontekście własnego działania. Może to bowiem rodzić tendencję do zatajania faktów kradzieży. Z drugiej jednak strony ujęcia eufemistyczne są z natury rzeczy mniej precyzyjne i jednoznaczne. Nie ma zatem pewności, czy pozytywna odpowiedź na pytanie o to, czy na przykład „zdarzyło ci się kiedyś zabrać komuś 500 zł?”, jest przez danego respondenta rozumiane jako słowny równoważnik pojęcia „kradzież”. Mógł on bowiem zabrać komuś i nie oddać tych 500 zł, lecz nie uważać tego działania za kradzież (bezwrotna pożyczka czy pożyczka „wymuszona”, depozyt, wyegzekwowanie długu itp.). Tak więc mimo że sformułowania eufemistyczne są zazwyczaj zbliżone do klasy zjawisk, o które pyta badacz, nigdy nie są z nimi tożsa-

me. Rodzi to niebezpieczeństwo, że część badanych zrozumie pytanie niezgodnie z intencjami pytającego.

Jednak pytania ujmowane wprost także nie muszą być jednoznacznie rozumiane. Trzymając się przykładu kradzieży: to co dla jednego respondenta będzie kradzieżą, dla innego będzie tylko „pożyczeniem sobie”, zabraniem rzeczy niczyjej, odebraniem rzeczy niesłusznie posiadanej przez inną osobę, odebraniem rzeczy własnej, wyrównaniem zobowiązań czy wyegzekwowaniem długu. Przykłady, które można by mnożyć, wskazują, że również w przypadku klarownego zdefiniowania sytuacji nie daje się uniknąć niebezpieczeństwa odmiennego od intencji badacza zrozumienia pytań kwestionariusza *self-report*.

Jak realna jest to groźba, niech świadczy przykład wyników badań Malewskiej i Muszyńskiego²⁶. Autorzy zmierzali m.in. do ustalenia, jakie zachowania są w oczach dzieci postrzegane jako kradzież. Okazało się, że dla wielu badanych takie czyny jak zabranie kredy ze szkoły, drobnego przedmiotu koledze, rzeczy stanowiącej mienie publiczne etc. w ogóle nie były postrzegane jako kradzieże. Odsetki badanej młodzieży, która nie uważała za akty kradzieży zachowań będących nimi w istocie, przekraczały niekiedy 50. Świadczy to, że pytając wprost o kradzież (a także np. o rozbój czy paserstwo) także nie daje się w pełni wyeliminować różnic w sposobie rozumienia zadawanych pytań.

Ponieważ zarówno pytania formułowane wprost, jak i eufemistycznie mogą nastroczać respondentom analogiczne kłopoty, jeśli chodzi o ich właściwe rozumienie, w badaniach *self-report* stosuje się raczej pytania ujmowane eufemistycznie, zgodnie z linią wytyczoną przez pionierskie prace Shorta i Nye'a²⁷. Autorzy ci są zdania, że w przypadku eufemistycznego określania charakteru czynu mniejsze jest niebezpieczeństwo, że respondenci będą mieć niemiłe dla nich skojarzenia, a co za tym idzie, mniejsza będzie tendencja do zatajania. Ten pogląd podziela zresztą wielu innych autorów²⁸.

Niemniej spotyka się też kwestionariusze *self-report*, w których niektóre przynajmniej pytania o zachowania dewiacyjne ujmowane są wprost²⁹. Wypada jednak zaznaczyć, że po pierwsze, brak jest systematycznych badań, w których stosowano by wariantowo pytania formułowane wprost i eufemistycznie, stąd nie jest bynajmniej oczywiste, że drugi sposób jest lepszy. Jak na razie opieramy się na intuicjach i przypuszczeniach. Po drugie, wiele czynów nie daje się operacjonalizować w postaci eufemistycznych pytań, np. pytania o udział w bójce czy pobiciu.

²⁶ H. Malewska, H. Muszyński: *op. cit.*

²⁷ J.F. Short, J. Nye: *op. cit.* (1957, 1958); J. Nye: *op. cit.*

²⁸ H. Muszyński: *op. cit.*; N. Christie i in.: *op. cit.*; D. Farrington: *op. cit.*; R. Hardt, S. Peterson-Hardt: *op. cit.*

²⁹ M. Hindelang: *op. cit.*; J. Shapland: *Self-reported Delinquency in Boys Aged 11 to 14*, „The British Journal of Criminology”, 1978, vol. 18.

Z powyższym łączy się ściśle kwestia stopnia abstrakcyjności ujmowanych w pytaniach sytuacji i definicji zachowań. Wspominaliśmy już, że obecnie zdecydowanie przeważa pogląd, iż zarówno czyny, jak i okoliczności ich dokonywania winny być raczej opisowe i eufemistyczne. Wyłania się jednak problem, jak ogólne mają być to opisy. Sięgnijmy do przykładu kradzieży. W badaniach *self-report* można pytać nader ogólnie: „czy zdarzyło ci się coś zabrać i nie oddać?” lub też konkretniej: „czy zdarzyło ci się coś zabrać i nie oddać ze sklepu?” lub jeszcze bardziej szczegółowo: „czy zdarzyło ci się zabrać i nie oddać ze sklepu cukierków?”. Różne mogą być przedziały wartości skradzionego mienia, które interesować będą badacza: kradzież rzeczy o wartości dwóch, pięciu, pięćdziesięciu dolarów itp. Różne mogą być wreszcie osoby „ofiary”: kradzież na szkodę kolegi, rodziców, kradzież w sklepie, kradzież osobie nieznamomej.

W kwestionariuszach zbliżonych do skali Shorta-Nye'a zamieszcza się różne przedziały wartości skradzionego mienia, lecz sytuacje, w których dokonano tych kradzieży, nie są wyspecyfikowane³⁰. Analogiczną technikę stosują badacze posługujący się kwestionariuszem Hirschiego³¹.

Natomiast w kwestionariuszach brytyjskich naukowców spotyka się najczęściej bardzo konkretne rodzaje zachowań oraz ich okoliczności. Przykładowo: w kwestionariuszu skonstruowanym przez Gibsona³², wykorzystywanym w wielu badaniach³³, pytania dotyczące kradzieży obejmowały

³⁰ Spośród licznych badań, w których wykorzystywano tę skalę, wskazać można przykładowo następujące: H. Voss: *Differential Association and Delinquent Behavior*, „Social Problems”, 1964, vol. 12; H. Voss: *Socio-economic Status and Reported Delinquent Behavior*, „Social Problems”, 1965, vol. 13; A. Reiss, A. Rhodes: *An Empirical Test of Differential Association Theory*, „Journal of Research in Crime and Delinquency”, 1964, vol. 1; V. Matthews: *Differential Identification. An Empirical Note*, „Social Problems”, 1968, vol. 15; S. Shokam, E. Ruth, W. Reckless: *Value Orientation and Awareness of Differential Opportunity of Delinquent and Non-delinquent Boys in Israel*; „Criminologica”, 1964, vol. 2; P. Friday: *The Applicability of Differential Opportunity and Differential Association in Sweden* (tekst powielony).

³¹ T. Hirschi: *op. cit.*; E. Linden, J. Hackler: *Affective Ties and Delinquency*, „Pacific Sociological Review”, 1973, vol. 27; G. Jensen: *Parents, Peers and Delinquent Action. A Test of Differential Association Perspective*, „American Sociological Review”, 1972, vol. 78; J. Hepburn: *Testing Alternative Models of Delinquent Causation*, „The Journal of Criminal Law and Criminology”, 1976, vol. 67.

³² H. Gibson: *Self-reported Delinquency among Schoolboys and Their Attitude to the Police*, „The British Journal of Social and Clinical Psychology”, 1967, vol. 6.

³³ H. Gibson, S. Morison, D. West: *The Confession of Known Offences in Response to a Self-reported Delinquency Schedule*, „The British Journal of Criminology”, 1970, vol. 10; D. Farrington: *op. cit.*; D. Farrington: *The Effect of Public Labelling*, „The British Journal of Criminology”, 1977, vol. 17; D. Farrington, S. Osborn, D. West: *The Persistence of Labelling Effects*, „The British Journal of Criminology”, 1978, vol. 18; J. Shapland: *op. cit.*

spory wachlarz sytuacji (kradzież z supermarketu, z małego sklepu, z domu, ze szkoły, z miejsca pracy itp.), jak również rozmaite wartości skradzionego mienia ³⁴. W kwestionariuszu tym uszczegółowione są także sytuacje dotyczące zachowań agresywnych (np. bójka z nieznajomymi, bójka z kolegami, bójka z policjantami).

Spotyka się również kwestionariusze, w których sytuacje i okoliczności czynów są niezmiernie precyzyjnie określone. I tak np. Ostrihanska i Wójcik ³⁵ ujmowały jako trzy odrębne rodzaje kradzieży: kradzież ze sklepu, kradzież cukierków ze sklepu oraz kradzież wina ze sklepu. Nietrudno zauważyć, że w takim przypadku niebezpieczeństwo bezzasadnego zawyżania faktycznych rozmiarów zachowań dewiacyjnych badanych jest szczególnie duże. Załóżmy bowiem, że staramy się oszacować liczbę kradzieży, których dopuszcza się konkretny respondent. Jeżeli ukradł on ze sklepu wino, to przyzna się do tego faktu, ale także przyzna, że w ogóle dopuścił się kradzieży ze sklepu, choć w istocie chodzi o ten sam czyn. Autorki wspomnianego badania nie przeprowadzały jednak takich oszacowań. Jeżeli chodzi natomiast ponadto o ustalenie, z jaką częstością dany czyn był dokonywany, wówczas błąd pomiaru może ulec dalszemu zwiększeniu. Przykładowo: skrupulatny respondent, który ukradł ze sklepu dwa razy cukierki i trzy razy wino udzieli na powyższe pytania następującej odpowiedzi: 2+3+5 (ponieważ dokonał ogółem pięciu kradzieży ze sklepu), co da w rezultacie liczbę dziesięciu kradzieży, podczas gdy w rzeczywistości było ich pięć.

Stosowanie tej procedury nie jest jednak tak bezzasadne, jak mogłoby to sugerować powyższy przykład. Wprawdzie z logicznego punktu widzenia operacja tego rodzaju wydaje się nieuprawniona (ponieważ klasy zjawisk są nierozłączne), jednak ma ona także pewne zalety. W rezultacie dużego uszczegółowienia sytuacji, w jakich nastąpiło dane zachowanie, wzrasta bowiem hipotetyczne prawdopodobieństwo, że badani przypomną je sobie. W pytaniach typu: „czy zdarzyło ci się kiedyś coś komuś zabrać i nie oddać?”, prawdopodobieństwo to jest teoretycznie mniejsze.

Ponadto w wyniku dużego skonkretyzowania sytuacji możemy starać się oszacować także wagę czynu, która wyraża się między innymi w wartości skradzionego mienia czy osobie ofiary. W analizowanym przykładzie chodziło zapewne o to, iż autorki przyjęły założenie, że kradzież ze sklepu wina jest jak gdyby odmienna jakościowo (poważniejsza) od kradzieży cukierków.

Generalnie rzecz biorąc kwestia stopnia abstrakcyjności pytań kwestionariuszy *self-report* winna być rozstrzygana przy uwzględnieniu celów, którym ma służyć to narzędzie badawcze. Wydaje się, że formułowanie bardzo szczegółowych pytań o zachowania dewiacyjne może okazać się ^{34 35}

³⁴ Por. także J. S h a p l a n d : *op. cit.*

³⁵ Z. O s t r i h a n s k a : *op. cit.*; Z. O s t r i h a n s k a, D. W ó j c i k : *op. cit.*

korzystne w badaniach, które mają na celu przede wszystkim względy poznawcze. Przykładowo: jeśli interesuje nas przede wszystkim struktura i rozmiary nie ujawnionych kradzieży, wówczas korzystne wydaje się pytanie o kradzieże z poszczególnych miejsc (małe sklepy, supersamy, miejsca pracy, piwnice) lub na szkodę poszczególnych podmiotów (osoby bliskie sprawcy, osoby nieznanym, instytucje, zakłady pracy). Otrzymujemy wówczas pełniejszy obraz rozpatrywanego zjawiska. Bardzo szczegółowych pytań należy natomiast unikać, gdy staramy się oszacować rozmiary dewiacji poszczególnych respondentów. Wówczas bowiem istnieje realne niebezpieczeństwo otrzymania mocno zniekształconego obrazu faktycznego poziomu zachowań dewiacyjnych poszczególnych badanych w rezultacie wielokrotnego zliczania jakościowo tożsamych zachowań.

W praktyce jednak większość kwestionariuszy *self-report* zawiera pytania o kradzież w rozmaitych przedziałach wartości, z rozmaitych miejsc czy na szkodę rozmaitych osób. Analogicznie postępuje się w przypadku pytań o stosowaniu agresji fizycznej. Najczęściej zadaje się odrębne pytania dotyczące pobicia kolegi, pobicia osoby nieznanym, pobicia policjanta, udziału w bójkę (ewentualnie z uwzględnieniem dalszego wyspecyfikowania sytuacji, w których dochodziło do bójki). Nie bezzasadne (choćby ze względu na kryteria formalnoprawne) wydaje się również wyodrębnienie sytuacji, w których stosowano agresję fizyczną, ze względu na stopień uszkodzenia ofiary (np. pobicie z obrażeniami), sposób działania (np. pobicie przy użyciu niebezpiecznych narzędzi) czy pobudki działania sprawcy (np. pobicie chuligańskie). Rzecz jasna, przy ocenie takiego podejścia nie można abstrahować od celów, którym służyć ma dany kwestionariusz. Jeżeli służy on jedynie jako punkt wyjścia do tworzenia skali, procedura polegająca na zamieszczeniu w nim nierozłączonych lub zazębiających się klas zjawisk nie nasuwa poważniejszych zastrzeżeń. Jeśli natomiast pytania kwestionariusza *self-report* już stanowią skalę (jak w przypadku skali Shorta-Nye'a), podejście takie nie wydaje się uprawnione. Skale typu guttmanowskiego nie wykluczają bowiem możliwości wielokrotnego zliczania tego samego zachowania.

Jednak wszystkie powyższe zastrzeżenia mają relatywnie mniejsze znaczenie, jeśli chodzi o oszacowanie globalnych rozmiarów dewiacyjnego zaangażowania konkretnego respondenta czy grup respondentów. Wówczas błędy pomiaru, o których była mowa, są zazwyczaj już z góry wkalkulowane, ponieważ zasadniczym celem tego typu kwestionariuszy jest wyodrębnienie grup różniących się pod względem poziomu dewiacji. Ponadto błąd, o którym mowa, daje się zazwyczaj korygować na dalszych etapach analizy np. w wyniku konstruowania skal dewiacji zawierających już znacznie mniejsze liczby zachowań, przez co maleje prawdopodobieństwo wielokrotnego zliczania tożsamych zakresowo klas zjawisk.

Na zakończenie rozważań poświęconych problematyce sposobów for-

mułowania pytań kwestionariuszy do badania nie ujawnionych zachowań dewiacyjnych wypada wspomnieć o generalnej dyrektywie badawczej, jaka winna obowiązywać przy opracowywaniu pytań o zachowania dewiacyjne. Ponieważ pewne niejasności i wieloznaczność zawsze kryją się w tych pytaniach, należy bezwzględnie dążyć do eliminowania z kwestionariuszy *self-report* wszelkich pytań, które już ze swej natury są wieloznaczne lub zawierają element oceny³⁶. Zaleca się np. unikanie pytań o takie czyny, jak nieostrożna jazda, sprzeniewierzenie lub gwałt. Pogląd ten jest w pełni uzasadniony. Skoro bowiem sąd jest często w dużym kłopotcie, jak zaklasyfikować takie zachowanie, tym większe więc problemy będą mieli ewentualni respondenci.

Z innych względów w kwestionariuszach *self-report* należy również unikać pytań o czyny bardzo rzadkie, np. zabójstwo. Istnieje bowiem bardzo małe prawdopodobieństwo znalezienia się w próbie dostatecznej do analizy statystycznej liczby osób, które dopuściły się takich czynów. Ta trafna dyrektywa metodologiczna nie zawsze jednak jest przestrzegana w praktyce. I tak np. Hindelang³⁷ zamieścił pytanie o zniszczenie mienia o wartości mniejszej i większej niż 10 dolarów. Jest wątpliwe, by (szczególnie młodzi) respondenci byli w stanie tak precyzyjnie oszacować wartość zniszczonych przez siebie rzeczy. W kwestionariuszu tym znalazły się także i inne, dość niefortunnie sformułowane pytania np. o „użycie siły w celu zwrócenia na siebie uwagi seksualnej dziewczyny”. Nie jest jasne, czy chodzi tu o gwałt, pogwizdywanie, czy podszcypywanie. Z kolei Farrington³⁸ posłużył się pytaniem o wyłudzenie pieniędzy, przy czym pojęcie to nie zostało poparte przykładem czy rozwinięte, a co za tym idzie, pozostawiało sporo swobody interpretacyjnej.

Autorzy skandynawscy także nie ustrzegli się tego niebezpieczeństwa³⁹, zamieszczając w kwestionariuszu pytanie o zakłócanie porządku publicznego w stanie nietrzeźwym. Wprawdzie jedna z możliwości odpowiedzi zakładała, że respondenci nie będą wiedzieli, czy ich zachowanie zakłócało porządek publiczny, to jednak pytanie to stanowi przecież klasyczny przypadek wieloznacznego i ocennego opisu sytuacji, których należy unikać w kwestionariuszach *self-report*.

3. LICZBA PYTAŃ

Powyżej analizowaliśmy najważniejsze zagadnienia, które pojawiają się w trakcie operacjonalizacji klas zachowań dewiacyjnych, wyrażające się w nadawaniu formy poszczególnym pytaniom kwestionariusza. Nie-

³⁶ Por. np. N. Christie i in.: *op. cit.*; D. Farrington: *op. cit.* (1973).

³⁷ M. Hindelang: *op. cit.* (1971).

³⁸ Farrington zwraca uwagę, że badani mieli nawet poważne kłopoty z właściwym rozumieniem zwrotu itp., który następował po egzemplifikacjach sytuacji.

³⁹ N. Christie i in.: *op. cit.*

mniej istotna jest kwestia samej liczby owych pytań. Jak podnoszą Christie⁴⁰ oraz Farrington⁴¹, uwzględnione w kwestionariuszu zachowania dewiacyjne stanowią zawsze jedynie pewną próbę takich zachowań. Przy tym jest to próba o nieznanym zazwyczaj parametrach. Dlatego też, wbrew opiniom zwolenników metody *self-report* formułowanym w początkowym okresie jej stosowania, nie jest bynajmniej łatwo przy jej pomocy oszacować przykładowo rozmiary „ciemnej liczby” przestępstw. Mamy tu bowiem do czynienia niejako z dwiema próbami: z próbą osób z jakiejś konkretnej, lecz zawsze ograniczonej w swych parametrach populacji (np. młodzież męska w wieku 18—24 lat) oraz z próbą zachowań przestępczych, przy czym i w tym przypadku nie wiemy zazwyczaj, jak reprezentatywna jest próba. Zatem możliwość błędu w oszacowaniu jest spora.

Obecnie jednak celem badań *self-report* nie jest zazwyczaj szacowanie rozmiarów nie ujawnionych zachowań dewiacyjnych w całych populacjach, lecz raczej ustalanie tych rozmiarów w rozpatrywanych zbiorowościach, co stanowi punkt wyjścia do dalszych analiz — np. do testowania teorii dewiacji. Także i w tym przypadku warto mieć jednak świadomość, że bierze się pod uwagę tylko wycinek wszystkich możliwych zachowań dewiacyjnych, a co za tym idzie, możliwość generalizacji jest ograniczona. Wydaje się, iż zakres generalizacji jest tym większy, im większy lub „lepszy” wycinek uda się nam uzyskać. Przykładowo: kwestionariusz *self-report* przeznaczony do badań ogólnokrajowych będzie obciążony poważnym błędem, jeśli uwzględni się w nim zachowania dewiacyjne, charakterystyczne wyłącznie dla populacji wielkomiejskich. Kwestionariusz *self-report* układany pod kątem zachowań dewiacyjnych chłopców będzie relatywnie mniej użyteczny do badania zachowań dewiacyjnych dziewcząt. Kwestionariusz *self-report* konstruowany pod kątem uczącej się młodzieży mieć będzie ograniczone zastosowanie do badania zachowań dewiacyjnych młodzieży pracującej etc.

Ogólnie rzecz biorąc sądzić można, że im więcej różnorodnych zachowań uwzględni dany kwestionariusz, tym bardziej „próba zachowań” staje się reprezentatywna w stosunku do „populacji zachowań dewiacyjnych”. Należy jednak zdawać sobie sprawę z faktu, iż odmienna najczęściej liczba uwzględnianych zachowań w różnych kwestionariuszach *self-report* ogranicza w pewnym stopniu możliwość porównywania rezultatów badań, jeśli chodzi o globalne rozmiary nie ujawnionej przestępczości i innych zachowań dewiacyjnych. Czym bowiem więcej pytań zawiera kwestionariusz, tym większe prawdopodobieństwo, że do któregoś z tych zachowań przyzna się dany respondent. Dlatego też szacunki nie ujawnionego poziomu przestępczości i innych zachowań dewiacyjnych różnią się

⁴⁰ *Ibidem.*

⁴¹ D. Farrington: *op. cit.*

nieraz dość znacznie, co — pomijając inne względy — bywa skutkiem odmiennej liczby uwzględnianych zachowań.

W światowej kryminologii i socjologii dewiacji można, zachowując duży stopień umowności, wyróżnić kilka „szkół”, jeśli chodzi o sposób rozwiązywania zagadnienia liczby pytań, co pozostaje w bezpośrednim związku z rodzajem kwestionariusza-matki, jaki stosuje dany badacz. Jak wspominaliśmy na wstępie, nadal najpopularniejszy jest kwestionariusz *self-report* opracowany przez Shorta i Nye'a, który stanowi skalogram Guttmana. Choć wyjściowa „próba pytań” Shorta i Nye'a obejmowała 23 pytania, właściwa skala *self-report* (w zależności od charakteru badanej zbiorowości) zawierała zazwyczaj około 10 pytań. Dlatego też badacze stosujący wspomnianą skalę mają stosunkowo skromną reprezentację zachowań dewiacyjnych⁴². Inny krótki kwestionariusz *self-report* opracował Hirsch⁴³. Obejmuje on sześć pozycji i jest obecnie największym konkurentem skali Shorta-Nye'a⁴⁴. Obydwa omawiane kwestionariusze mają tę wadę, że uwzględnia się w nich zbyt wiele czynów bardzo błahych. Ponadto można je stosować w dość specyficznych warunkach społeczno-kulturowych i w zasadzie wyłącznie do badania młodszej, uczącej się młodzieży.

Także Gold stworzył kwestionariusz obejmujący niewielką liczbę pytań⁴⁵. Nadaje się on wprawdzie także do ustalania czynów poważniejszych, lecz nie można go stosować w warunkach badania za pomocą ankiety audytoryjnej⁴⁶. W latach sześćdziesiątych pewną popularnością cieszył się również pięciopozycyjny kwestionariusz do badania rozmiarów kradzieży opracowany przez Dentlera i Monroe⁴⁷. Ponieważ w dalszym ciągu brak jest znormalizowanego narzędzia pomiaru do badań nie ujawnionych zachowań dewiacyjnych, wielu autorów nadal tworzy swe własne kwestionariusze. Z nowszych kwestionariuszy, które zawierają niewielką liczbę pytań, wskazać można przykładowo kwestionariusz opracowany przez Axenrotha, który obejmuje 12 pozycji⁴⁸.

Wprawdzie niewielkie kwestionariusze *self-report* uniemożliwiają pre-

⁴² M.in. J. F. Short: *Differential Association and Delinquency*, „Social Problems”, 1957, vol. 4 oraz idem: *Differential Association with Delinquent Friends and Delinquent Behavior*, „Pacific Sociological Review”, 1958, vol. 1 oraz *Differential Association as a Hypothesis: Problems of Empirical Testing*, „Social Problems”, 1960, vol. 8; H. Voss: *op. cit.* (1963, 1964); A. Reiss, A. Rhodes: *op. cit.*; S. Shokam i in.: *op. cit.*; V. Matthews: *op. cit.*; P. Friday: *op. cit.*

⁴³ T. Hirschi: *op. cit.*

⁴⁴ Kwestionariusz ten stosowali m.in. G. Jensen (*op. cit.*), E. Linden, J. Hackler (*op. cit.*), J. Hepburn (*op. cit.*).

⁴⁵ M. Gold: *op. cit.* (1963, 1966, 1970).

⁴⁶ Kwestionariusz opracowany przez Golda stosowała m.in. M. Morash: *angs, Groups and Delinquency*, „The British Journal of Criminology”, 1983, vol. 23.

⁴⁷ R. Dentier, L. Monroe: *op. cit.*

⁴⁸ Axenroth: *op. cit.*

czyjne ustalenie, ile razy dany respondent popełnił inne czyny niż uwzględnione w kwestionariuszu, nie oznacza to jednak, że kwestionariusz *self-report* z wąską listą zachowań jest gorszym narzędziem w szacowaniu poziomu dewiacji badanych. Wynika to z faktu, że wiadomo jest zazwyczaj, z jakimi innymi zachowaniami, nie uwzględnionymi w kwestionariuszu, zachowania uwzględnione są wysoko skorelowane. Dlatego też na podstawie relatywnie niewielu zachowań można dość precyzyjnie ustalać poziom dewiacji badanych.

Bardzo wiele zachowań dewiacyjnych zawierają natomiast brytyjskie kwestionariusze do badania nie ujawnionych zachowań dewiacyjnych⁴⁹. Co ciekawe, kwestionariusz Gibsona, który był następnie szeroko stosowany tak przez autora, współautorów, jak i innych badaczy, nadaje się doskonale do skrócenia czy ewentualnego nadania mu cech skali Guttmana. Fakt, że jest on nadal stosowany w nie zmienionej postaci, świadczyć by mógł o tym, iż badacze kierują się świadomą dyrektywą metodologiczną nakazującą uwzględnianie bardzo wielu różnorodnych zachowań dewiacyjnych młodzieży (kwestionariusz ten, w zależności od wersji, liczy około 40 pytań).

W dalszym ciągu nie można mówić więc o jakiejś dominującej tendencji, jeśli idzie o liczbę zachowań uwzględnianych w kwestionariuszach *self-report*. Spośród innych kwestionariuszy, w których uwzględniano stosunkowo dużą liczbę zachowań, wskazać można przykładowo następujące: Christie i in.⁵⁰ uwzględnili 25 czynów, Clark i Wenninger⁵¹ — 38 czynów, Hardt zaś i Peterson-Hardt⁵² — 21 czynów. Wydaje się, że bezpieczniej jest stosować kwestionariusze obejmujące większe liczby zachowań dewiacyjnych. Daje to również większe możliwości dokonywania porównań w czasie czy między różnymi populacjami, choć wypełnianie takiej ankiety jest bardziej czasochłonne.

4. CEZURA CZASOWA

W badaniach metodą *self-report* spotyka się najczęściej dwa sposoby podejścia do zagadnienia przedziału czasu między dokonaniem czynu a chwilą, w której przeprowadzane jest badanie.

Pierwszy sposób polega na niewprowadzaniu cezury czasowej. Wówczas pytanie ma zazwyczaj formę: „czy kiedykolwiek zdarzyło ci się...”.

⁴⁹ Zob. m.in. H. Gibson: *op. cit.*; H. Gibson, S. Moris on, D. West: *op. cit.*; D. Farrington: *op. cit.* (1973, 1977); J. Shapland: *op. cit.*; H. Wilson: *Parental Supervision: a Neglected Aspect of Delinquency*, „The British Journal of Criminology”, 1980, vol. 18.

⁵⁰ N. Christie i in.: *op. cit.*

⁵¹ J. Clark, E. Wenninger: *Socio-economic Class and Area as Correlates of Illegal Behavior among Juveniles*, „American Sociological Review”, 1962, vol. 27.

⁵² R. Hardt, S. Peterson - Hardt: *op. cit.*

Badacz, który wybiera takie rozwiązanie, zmierza do wychwycenia wszystkich wyszczególnionych w kwestionariuszu zachowań lub czynów bez względu na to, jak dawno się zdarzyły. W swych pionierskich badaniach Short i Nye tak właśnie postąpili⁵³. Dlatego też większość badań pochodzących z końca lat pięćdziesiątych i z lat sześćdziesiątych, w których wykonywano skalę Shorta-Nye'a⁵⁴, nie zawierała ograniczeń czasowych⁵⁵.

Zwolennikiem odmiennego podejścia jest Hirschi⁵⁶. Autor ten jest zdania, że brak określenia cezury czasowej stanowił błąd wcześniejszych badań typu *self-report*, ponieważ nie można w ten sposób kontrolować związku wieku z zachowaniem dewiacyjnym. Wniosek swój opiera Hirschi na następującym rozumowaniu. Ponieważ starsza młodzież ma więcej okazji, by dokonać każdego z czynów uwzględnionych w kwestionariuszu, a także więcej okazji, by dokonać ich częściej, brak cezury czasowej powoduje niejako automatycznie, że starsza młodzież uzyskuje znacznie wyższe rezultaty na skalach dewiacji. Zaproponowane przez Hirschiego rozwiązanie miało w założeniach eliminować czynnik wieku badanych w wyniku zakreślenia rocznej cezury czasowej między chwilą dokonania danego czynu a chwilą przeprowadzania badania. Propozycja Hirschiego spotkała się z uznaniem, o czym świadczy fakt, iż obecnie przyjmuje się już niemal powszechnie roczną cezurę czasową w badaniach za pomocą metody *self-report*⁵⁷.

Decyzja o przyjęciu jednego lub drugiego rozwiązania zależy przede wszystkim od celów, którym ma służyć badanie *self-report*. Jeżeli w badaniu przeważają względy heurystyczne, innymi słowy, jeżeli badacz jest przede wszystkim zainteresowany w ustaleniu rozmiarów zjawiska w rozpatrywanej populacji, wówczas właściwsze wydaje się niewprowadzanie ograniczeń czasowych. Nie traci się wówczas z pola widzenia całości zjawiska i nie ogranicza się dość sztucznie faktycznego poziomu lub nasilenia dewiacji w analizowanej zbiorowości. Jeżeli natomiast kwestionariusz

⁵³ Mówiąc ściślej, Short i Nye pytali o czyny dokonane z chwilą rozpoczęcia nauki szkolnej. Gdy jednak zważyć, że badali oni młodzież w wieku 15—17 lat, oznaczało to w praktyce, jak już wspominaliśmy, rezygnację z cezury czasowej.

⁵⁴ M.in. H. Voss: *op. cit.* (1963, 1964); A. Reiss, A. Rhodes: *op. cit.* (1966).

⁵⁵ Jednym z nielicznych badań z początku lat sześćdziesiątych, w którym zastosowano roczną cezurę czasową, było badanie Clarka i Wenningera (*op. cit.*).

⁵⁶ T. Hirschi: *op. cit.*

⁵⁷ Por. np. M. Hindelang: *op. cit.* (1971, 1976); L. Biron, M. Leblanc: *Family Components and Home-based Delinquency*, „The British Journal of Criminology”, 1977, vol. 17; R. Simons, M. Miller, S. Aigner: *Contemporary Theories of Deviance and Female Delinquency: an Empirical Test*, „Journal of Research in Crime and Delinquency”, 1980, vol. 17; W. Lipton, D. Smith: *Explaining Delinquent Involvement: a Consideration of Suppressor Effects*, „Journal of Research in Crime and Delinquency”, 1983, vol. 20; W. Feyerherm: *The Group Hazard Hypothesis: a Reexamination*, „Journal of Research in Crime and Delinquency”, 1980, vol. 17; Axenroth: *op. cit.*

self-report, służy jedynie jako narzędzie pomiaru, jest środkiem do osiągnięcia jakiegoś innego celu wyznaczonego przez badacza ⁵⁸, to wówczas istotnie korzystne może okazać się określenie dokładnej cezury czasowej, np. rocznej. Przyjęcie takiego rozwiązania może bowiem zmniejszyć wpływ wieku badanych na zaobserwowane rozmiary dewiacji. Ale czy rzeczywiście udaje się w ten sposób całkowicie kontrolować zmienną wieku, jak twierdzi Hirschi? Wydaje się, że nie. Przystępczość bowiem czy, szerzej, dewiacja nie stanowi naszym zdaniem zwykłej funkcji czasu w znaczeniu astronomicznym. Chodzi tu raczej o czas (wiek) w znaczeniu biologicznym, fizjologicznym i psychicznym. Prawdą jest, że rozmiary dewiacji wzrastają z wiekiem. Jednak zależność ta nie wypływa wyłącznie z faktu, że starsza młodzież miała uprzednio więcej okazji, by dopuścić się inkryminowanych zachowań. Nie mniej istotne jest i to, że pewne zachowania zaczynają się pojawiać bądź nasilać dopiero w starszym wieku. Dlatego też określenie sztywnej cezury czasowej i tak nie prowadzi do całkowitego eliminowania czynnika wieku, powoduje natomiast utratę wielu cennych informacji o faktycznych rozmiarach dewiacji w analizowanych zbiorowościach.

Przykładowo: Hardt i Peterson-Hardt⁵⁹ porównywali rozkłady pozytywnych odpowiedzi na pytania o zachowania dewiacyjne w dwóch wariantach — bez cezury czasowej i z roczną cezurą czasową. Zgodnie z oczekiwaniami okazało się, że odsetki pozytywnych odpowiedzi były w pierwszym wariacie znacznie wyższe niż w wariacie drugim ⁶⁰.

Spotyka się także kwestionariusze *self-report*, w których stosuje się wariant pośredni. Polega on na tym, że respondenci mają za zadanie określić, ile razy dopuszczali się danego czynu w ciągu ostatniego roku (miesiąca, tygodnia) oraz ile razy popełnili dany czyn dawniej niż rok temu ⁶¹.

Kwestionariusze *self-report* bez cezury czasowej istotnie wymagają zwrócenia bacniejszej uwagi na czynnik wieku ⁶². Mają one jednak tę zaletę, że za ich pomocą otrzymujemy pełniejszy obraz poziomu dewiacji w rozpatrywanych zbiorowościach i pozbawiony zniekształceń, wynikających z wprowadzenia ograniczeń czasowych. Natomiast czynnik wieku i tak można kontrolować za pomocą stosunkowo prostych metod statystycznych (np. analiza regresji).

⁵⁸ Obecnie najczęściej na podstawie wyników uzyskiwanych przez respondentów na skalach *self-report* konstruuje się zmienną zależną „zachowanie dewiacyjne”, której związki ze zmiennymi niezależnymi uwzględnionymi przez badacza są następnie analizowane.

⁵⁹ R. Hardt, S. Peterson - Hardt: op. cit.

⁶⁰ Analogiczne rezultaty uzyskał uprzednio D. Farrington: op. cit. (1973).

⁶¹ M.in. T. Hirschi: op. cit.; D. Farrington: op. cit. (1973); R. Hardt, S. Peterson - Hardt: op. cit.; W. Lipton, D. Smith: op. cit.

⁶² Jest to szczególnie istotne w badaniach, których zasadniczym celem jest określenie zmiennych predykcyjnych względem zachowania dewiacyjnego, ponieważ niekontrolowanie zmiennej wieku może doprowadzić do uznania za rzeczywistą zależność pozornej, wypływającej z wieku.

Jednak za wprowadzaniem cezury czasowej także przemawiają istotne argumenty. Pomijając już analizowany uprzednio problem różnic w wieku badanych, określenie czasu poprzedzającego badanie, do którego winny odnosić się ujawniane w kwestionariuszu zachowania, ma dwie inne zalety. Po pierwsze, eliminujemy w ten sposób kwestię tych respondentów, którzy, przypuścmy, w wieku 12—14 lat dopuszczali się bardzo wielu zachowań wyszczególnionych w kwestionariuszu, jednak w momencie realizacji badania wydorośleli już i spowaźnieli, w związku z czym ich aktualny poziom dewiacji jest relatywnie niski. Brak cezury czasowej uniemożliwia uchwycenie tego rodzaju zmian w zachowaniu. Po drugie, wprowadzenie cezury czasowej, szczególnie krótkiej, rocznej lub jeszcze krótszej, umożliwia zminimalizowanie problemu zawodności pamięci badanych. Z pewnością łatwiej przypomnieć sobie np. fakt niszczenia przedmiotów publicznego użytku, jeżeli nastąpił on w ciągu roku poprzedzającego badanie, aniżeli 5 czy 7 lat wcześniej.

Konkludując uwagi dotyczące zagadnienia cezury czasowej godzi się wspomnieć, że jakkolwiek dwa omawiane powyżej podejścia są obecnie dominujące (tj. roczna cezura czasowa lub całkowite rezygnowanie z jej określania), to niekiedy spotyka się też i inne. I tak np. Simson i in. wyznaczyli dwumiesięczną cezurę czasową⁶³, Lipton i Smith zaś cezurę osiemnastomiesięczną⁶⁴. Trudno mówić więc o zgodności stanowisk w tej istotnej kwestii.

5. TYPY ODPOWIEDZI

Decyzja o wyborze jednego z rozpatrywanych powyżej rozwiązań implikuje także zazwyczaj określony sposób budowy zestawu odpowiedzi. O ile jeszcze po wyznaczeniu rocznej lub krótszej cezury czasowej można starać się o stosowanie enumeratywnego zestawu odpowiedzi (co i tak może się okazać nieuzasadnione — np. w stosunku do bardzo błahych i rozpowszechnionych zachowań, jak chociażby jazda „na gapę”), to już rezygnowanie z zakreślania cezury czasowej czyni z enumeratywnego zestawu odpowiedzi wymóg całkowicie nierealny. W większości przypadków bowiem trudno oczekiwać, aby badani pamiętali dokładnie, ile razy dopuścili się na przykład kradzieży pieniędzy rodzicom w ciągu całego swego życia, tym bardziej gdy zdarzało się to dość często.

Dlatego też, analogicznie jak w przypadku zagadnienia cezury czasowej, kwestia zestawu odpowiedzi bywa najczęściej rozstrzygana dwojako. Enumeratywne zestawy odpowiedzi (tzn. „nigdy”, „jeden raz”, „dwa razy”, „trzy razy”, „cztery razy” itd.) dominują w kwestionariuszach typu *self-report*, w których wyznacza się krótką cezurę czasową. Natomiast

⁶³ R. Simson i in.: *op. cit.*

⁶⁴ W. Lipton, D. Smith: *op. cit.*

mniej precyzyjne typy odpowiedzi przeważają w kwestionariuszach tych autorów, którzy zrezygnowali z wyznaczania cezurę czasowej. W kwestionariuszach tych operuje się określeniami w rodzaju: „bardzo rzadko”, „często”, „bardzo często” lub też: „raz czy dwa razy”, „kilka razy”, „kilkanaście razy”⁶⁵.

Z pozoru wydawać by się mogło, że rozwiązanie enumeratywne jest lepsze, bo bardziej precyzyjne. W rzeczywistości jednak jego przyjęcie zakłada najczęściej zbyt wygórowane oczekiwania w stosunku do pamięci badanych. Czy istotnie bowiem młody człowiek jest w stanie udzielić tak dokładnej odpowiedzi na pytanie o częstość bójek z kolegami czy też częstość, z jaką chodził na wagary? Z pewnością w grupie respondentów, których cechuje bardzo niski poziom dewiacji, udzielenie precyzyjnej o powieździ nie będzie sprawiało specjalnego problemu. Przeważać tu będzie zdecydowanie odpowiedź typu „nigdy”. Również wśród badanych, którzy wykazują bardzo wysoki poziom dewiacji, udzielenie precyzyjnej odpowiedzi na większość pytań z enumeratywnym zestawem odpowiedzi nie będzie sprawiało trudności. Przeważać będzie tu typ odpowiedzi w rodzaju: „cztery razy lub częściej” (o ile kwestionariusz przewiduje taki rodzaj odpowiedzi, co nie jest regułą). Jednak dla badanych o średnim poziomie dewiacji udzielenie dokładnej odpowiedzi na pytanie: „ile razy” dopuścili się konkretnego czynu, nie będzie łatwe, niezależnie od tendencji do zatajania lub przesadzania. Po prostu najczęściej bardzo trudno jest przypomnieć sobie, ile razy nastąpiło inkryminowane wydarzenie. W sytuacji tego rodzaju z reguły nasuwa się odpowiedź w rodzaju: „może raz, może dwa razy”, której ten typ zestawu nie uwzględnia.

Niezależnie od powyższych wątpliwości można żywić uzasadnione podejrzenia, że badacze stosujący enumeratywne zestawy usiłują w ten sposób przejść na silniejszy, interwałowy typ skali. W rzeczywistości jednak zarówno zestaw enumeratywny, jak i zestawy mniej precyzyjne są w istocie skalami porządkowymi. Dlatego też tendencja do nadawania zestawowi enumeratywnemu cech skali interwałowej nie jest poprawna metodologicznie.

W praktyce jednak także i typowym skalom porządkowym (tj. zestawom odpowiedzi w rodzaju: „raz czy dwa razy”, „kilka razy”, „kilkanaście razy” itp.) nada je się cechy skali interwałowej⁶⁶. Poszczególne ty-

⁶⁵ Np. H. Muszyński: *op. cit.*; E. Vaz: *Self-reported Delinquency and Youth Culture Activity*, „The Canadian Review of Sociology and Anthropology”, 1965, vol. 2; D. Farrington: *op. cit.* (1973); R. Simons i in.: *op. cit.*; Axenroth: *op. cit.*; W. Feyerherm: *op. cit.*

⁶⁶ Por. np. J. Figueira-McDonough: *Feminism and Delinquency*, „The British Journal of Criminology”, 1984, vol. 24. Stosowany przez autorkę kwestionariusz *self-report* jest o tyle interesujący, że mimo wprowadzenia rocznej cezurę czasowej zamieszczono w nim dość ogólnikowe typy odpowiedzi: „raz czy dwa razy w roku”, „raz co dwa lub trzy miesiące”, „raz w ciągu każdych dwóch lub trzech tygodni”, „raz w tygodniu”, „dwa lub trzy razy tygodniowo”.

pom odpowiedzi przypisuje się bowiem dla celów analizy statystycznej odpowiednie wartości punktowe. Przykładowo: odpowiedź „nigdy” — 0 punktów, „raz czy dwa razy” — 1 punkt, „kilka razy” — 2 punkty, „kilkanaście razy” — 3 punkty itd.

Jak więc widać, omawiane sposoby konstruowania zestawów odpowiedzi różnią się w większej mierze odmiennymi założeniami metodologicznymi ⁶⁷, a także intuicjami badacza co do zalet określonej metody, niż praktycznymi implikacjami statystycznymi. Podobnie jak wiele innych zagadnień metodologicznych w badaniach typu *self-report* także i to nie zostało jeszcze definitywnie rozstrzygnięte, a co za tym idzie, decyzja o wyborze takiego czy innego sposobu konstrukcji zestawu odpowiedzi bywa najczęściej dość arbitralna. Na rzecz wariantu enumeratywnego przemawia to, że jego całkowite przeciwieństwo (tj. bardzo nieprecyzyjne rodzaje odpowiedzi, jak: „często”, „rzadko” ⁶⁸) nasuwa poważne obawy o właściwe i jednoznaczne rozumienie przez badanych takich ogólnikowych i subiektywnych ze swej natury określeń częstości. To co w odczuciu jednego respondenta będzie „częste”, dla innego może się okazać „rzadkie”. Z kolei jednak „ściśle enumeratywne” zestawy odpowiedzi (tj. zbudowane na zasadzie, „jeden raz”, „dwa razy” itp.) świadczą o tym, że ich autorzy mają chyba zbyt optymistyczne oczekiwania co do pamięci i dokładności respondentów.

Decyzja o wyborze rodzaju zestawu odpowiedzi nie jest zatem łatwa, a z metodologicznego punktu widzenia istotna. Wydaje się, że przy jej podejmowaniu winno się uwzględniać przede wszystkim przyjęte wcześniej rozwiązanie w kwestii cezur czasu oraz założony stopień dokładności pomiaru.

6. KOLEJNOŚĆ ODPOWIEDZI

Niezależnie od rodzaju zestawu odpowiedzi w badaniach typu *self-report* znaczenie mieć może także ich kolejność. Short i Nye ⁶⁹ wychodzili z założenia, że odpowiedzi na pytania kwestionariusza *self-report* powinny być zamieszczane w przemiennej kolejności, aby wyeliminować tendencję do stereotypowych czy nawykowych odpowiedzi.

Dajmy przykład. Jeżeli w pytaniu A zestaw odpowiedzi zaczyna się od „nigdy”, a kończy na „bardzo często”, to w pytaniu B powinien on być, zdaniem Shorta i Nye’a, odwrotny — od „bardzo często” do „nigdy”. W wy-

⁶⁷ Innymi słowy, chodzi tu o sposoby odpowiedzi na pytanie, która z następujących metod jest lepsza: czy pytać respondentów o dokładną liczbę czynów narażając się tym samym na błąd pomiaru, czy też już z góry założyć określony błąd pomiaru przez konstrukcję mniej precyzyjnych zestawów odpowiedzi.

Ten rodzaj zestawów odpowiedzi zawiera kwestionariusz opracowany przez Gibsona (op. cit.).

⁶⁹ J. Nye, J. F. Short: op. cit.; J. F. Short, J. Nye: op. cit.; J. Nye: op. cit.

niku zastosowania tej procedury utrzymuje się badanych w stanie „podwyższonej gotowości psychicznej”, ponieważ zmienna kolejność typów odpowiedzi zmusza niejako każdorazowo do zastanowienia się. Propozycja Shorta i Nye’a ma także ten walor, że już na wstępie analizy danych umożliwia w pewnym stopniu wyeliminowanie tych respondentów, którzy nie potraktowali poważnie postawionego przed nimi zadania, niezależnie od bardziej skomplikowanych metod ustalania szczerości wypowiedzi badanych. Dokonuje się tego na podstawie porównywania pierwszych i ostatnich rodzajów odpowiedzi na każde z pytań. Jeżeli bowiem dany respondent udziela na wszystkie pytania kwestionariusza odpowiedzi pierwszego typu, niezależnie od tego, czy jest nią akurat odpowiedź „nigdy”, czy „bardzo często”, możemy żywić uzasadnione obawy co do rzetelności jego odpowiedzi. Jest bowiem wysoce nieprawdopodobne, by, przypuśćmy, pięciu z dziesięciu czynów uwzględnionych w kwestionariuszu dopuszczał się on „bardzo często”, w ogóle zaś nie dopuszczał się pięciu pozostałych, choć i takiej możliwości nie można całkowicie wykluczyć.

Omawiany sposób sytuowania typów odpowiedzi ma jednak tę wadę, że wydłuża czas badania (przemienne odpowiedzi są trudniejsze do zrozumienia i wymagają większego wysiłku intelektualnego), co nie jest bez znaczenia w sytuacji, gdy badacz ma ograniczony czas na przeprowadzenie badania (np. w czasie jednej godziny lekcyjnej).

Jednak dyrektywa metodologiczna Shorta i Nye’a, mimo jej niewątpliwych zalet, nie została bynajmniej powszechnie zaakceptowana. Choć większość autorów zdaje się traktować problem kolejności odpowiedzi jako marginalny i nie zasługujący na osobną wzmiankę przy okazji prezentacji metodologicznej strony badań, to jednak lektura wielu prac skłania do przypuszczenia, że wymóg powyższy często nie był spełniany.

Przykładowo: Hirschi, który stworzył odrębną „szkołę” badań typu *self-report*, nie zastosował przemianych typów odpowiedzi. Przypuszczać można, że również w wielu innych badaniach, w których wykorzystywano kwestionariusz Hirschiego, nie były stosowane odpowiedzi przemienne. Natomiast w badaniach, w których stosowano skalę Shorta i Nye’a, metoda ta była zazwyczaj stosowana⁷⁰.

Jednak niezależnie od istniejącej w tej mierze praktyki badawczej autorzy zajmujący się metodologiczną problematyką badań *self-report* zalecają na ogół stosowanie przemiennej kolejności odpowiedzi na pytania o zachowania dewiacyjne⁷¹.

⁷⁰ Z nowszych badań, przy referowaniu których mówi się, że zestawy odpowiedzi zamieszczane były w przemiennej kolejności, wskazać można przykładowo tekst J. F i g u e i r a - M c D o n o u g h: *op. cit.*

⁷¹ D. F a r r i n g t o n: *op. cit.* (1973); R. H a r d t, S. P e t e r s o n - H a r d t: *bp. cit.*

Badacz przystępujący do badań typu *self-report* musi rozstrzygnąć kwestię, czy pytania dotyczące zachowań dewiacyjnych mają stanowić jeden blok (lub zgoła odrębny kwestionariusz), czy też winne one stanowić część jakiejś większej całości — np. „kwestionariusza do badań opinii młodzieży”. Nawet w przypadku gdy zdecydowano, iż kwestionariusz *self-report* będzie stanowił odrębną całość, otwarta pozostaje kwestia, czy mają się w nim znajdować wyłącznie pytania o zachowania dewiacyjne, czy też należy przeplatać je pytaniami o charakterze neutralnym lub pozytywnym („dobre uczynki”).

Intuicja podpowiada, że lepiej jest, gdy pytania dotyczące zachowań dewiacyjnych stanowią wycinek jakiejś większej całości, kiedy są przemieszane z innymi pytaniami: o wartości, postawy, opinie etc. Wówczas bowiem mniejsza jest podejrzliwość respondentów i w pewnej mierze można starać się zataić rzeczywisty charakter i cel badania. Ponadto zastosowanie rozwiązania, które polega na „maskowaniu” pytań o zachowania dewiacyjne, pozwala zmniejszyć lub wyeliminować tendencję do stereotypowych, nawykowych odpowiedzi, co sygnalizowaliśmy już powyżej omawiając kwestię przemiennej kolejności typów odpowiedzi.

Ponieważ obecnie już powszechnie technikę *self-report* stosuje się nie ze względu na próby oszacowania rozmiarów nie ujawnionych zachowań dewiacyjnych, a więc niesamoistnie, w przytłaczającej większości badań pytania o zachowania dewiacyjne stanowią część szerszych kwestionariuszy. Chodzi bowiem o to głównie, by w wyniku zastosowania tej techniki móc skonstruować zmienną zależną „zachowanie dewiacyjne” służącą dalszym analizom. Spotyka się jednak nadal badania, w których już to do kwestionariusza *self-report* dołączone są tylko pytania metryczkowe⁷², już to pytania o zachowania dewiacyjne, mimo że stanowią fragment większego kwestionariusza, wyodrębniane są graficznie lub nawet rozdawane są badanym na odrębnych kartkach⁷³. W tych przypadkach jednak problem wywołania specyficznego nastawienia badanych czy tendencji do stereotypowych odpowiedzi rysuje się nadal z całą ostrością⁷⁴.

Pomimo że zasadniczym celem wielu dotychczasowych badań typu *self-report* było konstruowanie prostego i krótkiego narzędzia pomiaru rozmiarów czy nasilenia zachowań dewiacyjnych w populacjach „porządnej” młodzieży, zasadne wydaje się zamieszczanie nawet w bardzo krótkich kwestionariuszach pytań o charakterze neutralnym czy pozytywnym⁷⁵ lub pytań o to, czy badani sami byli ofiarami przestępstw⁷⁶.

⁷² Np. Axenroth: *op. cit.*

⁷³ Np. T. Hirschi: *op. cit.*

⁷⁴ Szerzej ten temat zob. np. R. Hardt, S. Peterson-Hardt: *op. cit.*

⁷⁵ Np. D. Farrington: *op. cit.* (1973); R. Hardt, S. Peterson-Hardt: *op. cit.*

⁷⁶ N. Christie i in.: *op. cit.*

Wprawdzie sprawia to, że narzędzie pomiaru staje się trochę mniej eleganckie, a przede wszystkim wymaga przy zastosowaniu więcej czasu i jest większe objętościowo, jednak zalety sposobu, który polega na przeplataniu pytań o zachowania dewiacyjne innymi pytaniami, są, jak się wydaje, bezsporne.

7. ZAGADNIENIE WAG

Istotny problem teoretyczny i metodologiczny wyłania się z chwilą, gdy na podstawie odpowiedzi na pytania kwestionariusza *self-report* usiłujemy określić poziom czy też „głębnię” dewiacyjnego zaangażowania każdego z respondentów lub grup respondentów wyłonionych na podstawie wyników uzyskiwanych na skalach dewiacji. W kwestionariuszu *self-report* znajdują się zazwyczaj czyny o bardzo różnym „ciążarze gatunkowym”. Obok drobnych nieuczciwości i przejawów tzw. złego zachowania kwestionariusze *self-report* zawierają także pytania o czyny, których „ładunek dewiacji” jest znacznie większy (np. zażywanie narkotyków lub przestępstwa o poważnym charakterze, jak rozbój z bronią w ręku, ciężkie uszkodzenie ciała czy włamanie). Wyłania się zatem problem, czy wszystkie te zachowania, których waga jest przecież bardzo różna, traktować jednakowo, czy też w jakiś sposób je różnicować w zależności od „ładunku dewiacji”, jaki zawierają. Jeśli zaś uwzględniane w kwestionariuszach *self-report* zachowania różnicować przez nadawanie im odmiennych wag, to jakie kryteria wówczas stosować. Czy mają być to kryteria czysto formalne, np. przez nadawanie czynom o cechach wykroczenia wagi jeden, o cechach występku — wagi dwa, czynom zaś o znamionach zbrodni — wagi trzy? Czy o nadawaniu różnych wartości punktowych poszczególnym zachowaniom mają decydować tzw. sędziowie kompetentni? Czy też może rangować uwzględnione w kwestionariuszu zachowania w sposób arbitralny? Lub może w ogóle zrezygnować z przyznawania wag poszczególnym rodzajom zachowań traktując je wszystkie jednakowo? Są to kwestie bardzo ważne, ponieważ od sposobu ich rozstrzygnięcia zależy często, jaka grupa badanych będzie stanowiła grupę badawczą (tj. „nie ujawnionych przestępców”), a jaka grupę kontrolną (tj. „nieprzestępców”).

W przeważającej części badań typu *self-report* wagi nie są stosowane, tj. wszystkim uwzględnianym zachowaniom nadaje się analogiczną wartość punktową, uznając je za dewiacyjne w tym samym stopniu. Rozwiązanie takie jest jednak niedoskonałe, ponieważ intuicyjnie choćby wiadomo, że inny jest „ładunek dewiacji”, przypuśćmy, „jazdy na gape”, inny zaś rozboju. Stąd czynione były próby określenia powagi przestępstw i innych zachowań dewiacyjnych.

I tak np. Wolfgang i Sellin⁷⁷ dążyli do stworzenia znormalizowanej

⁷⁷ T. Sellin, M. Wolfgang: *Measurement of Delinquency*, New York 1964.

skali wagi przestępstw. Na podstawie oceny czynów, dokonywanej na dziesięciopunktowej skali przez respondentów, stworzyli oni skalę do ustalenia wagi przestępstw, która była następnie wykorzystywana m.in. do rangowania zachowań w badaniach typu *self-report*⁷⁸.

Przyjrzyjmy się, jak wygląda zastosowanie skali Wolfganga-Sellina w badaniach typu *self-report*. Oto Morash⁷⁹ w trakcie wywiadów kwestionariuszowych ustalała najpierw (procedura sortowania kart przez respondentów) pozytywne odpowiedzi na pytania o zachowania przestępcze. Następnie zaś nadawała wagi tym zachowaniom posługując się skalą Wolfganga-Sellina. Każde zachowanie uzyskiwało wagę w zależności od wartości skradzionego bądź uszkodzonego mienia, charakteru zajścia i stopnia pokrzywdzenia ofiary. Na przykład zniszczenie mienia o wartości poniżej 5 dolarów uzyskiwało wagę 1, powyżej 5 dolarów — wagę 2. Celowe uszkodzenie ciała otrzymywało wagę 1, kiedy było niewielkie lub gdy „stosowano narzędzia, nie przeznaczone zasadniczo do bicia”, wagę 4, kiedy nastąpiło poważne uszkodzenie ciała lub była używana broń, w przypadku zaś, gdy ofiara zajścia zmarła lub wymagała hospitalizacji oraz była używana broń — wagę 7. Groźba karalna miała w badaniu Morash wagę 2. Gdy jednak grozono „pobiciem pięściami lub nogami”, waga wzrastała do 4, kiedy zaś grozono użyciem broni — waga wynosiła 7. Zabór pojazdu w celu krótkotrwałego użycia uzyskiwał wagę 2 bez względu na okoliczności. Autorka nadawała wagi kradzieżom od 1 do 6 w zależności od skradzionej kwoty i okoliczności towarzyszących (np. czy stosowano siłę lub grozono bronią). Przykładowo: wagę 1 miała kradzież poniżej 5 dolarów, wagę zaś 6 — rozbój przy użyciu broni. Następnie Morash tworzyła indeks dewiacji obejmujący częstość (maksymalnie dwa razy i z roczną cezurą czasową), różnorodność i wagę czynów⁸⁰.

⁷⁸ Jednak zarówno metoda badawcza Sellina i Wolfganga, jak też stworzona przez nich skala do ustalania wagi przestępstw spotkały się z krytyką. Zob. G. Rose: *Concerning the Measurement of Delinquency*, „The British Journal of Criminology”, 1966, vol. 6; H. Leisure, P. Lehman: *Remeasuring Delinquency: a Replication and Critique*, „The British Journal of Criminology”, 1975, vol. 15; H. Wagner, K. Pease: *On Adding up Scores of Offence Seriousness*, „The British Journal of Criminology”, 1978, vol. 18; J. Sheley: *Crime Seriousness Rating*, „The British Journal of Criminology”, 1980, vol. 20; S. Gottfredson, K. Young, W. Laufer: *Additivity and Interactions in Offence Seriousness Scales*, „Journal of Research in Crime and Delinquency”, 1980, vol. 17. Podkreśla się przede wszystkim, że skala ta traktowana jest jako interwałowa lub zgoła ilorazowa, podczas gdy w istocie ma ona jedynie cechy skali porządkowej, choć Leisure i Lehman (op. cit.) kwestionują nawet porządkowy charakter skali Sellina i Wolfganga. Dlatego też nieuprawnione jest dokonywanie na tej skali operacji typu mnożenia, dzielenia czy dodawania, jak sugerowali to jej twórcy. Szersze rozważania nad tą metodą określania powagi przestępstw wykraczają już jednak poza ramy niniejszego opracowania.

⁷⁹ M. Morash: op. cit.

⁸⁰ *Ibidem*.

Inna metoda polega na wykorzystywaniu opinii tzw. sędziów kompetentnych. Na przykład Christie i in.⁸¹ wzięli uwzględnione w badaniu zachowania dewiacyjne pod kątem ładunku dewiacji na podstawie opinii (rangowania) dwóch sędziów — profesorów prawa karnego. Specjaliści ci mieli za zadanie uporządkować listę zachowań dewiacyjnych na dziesięciopunktowej skali. Między dokonanymi uporządkowaniami zachodził dość wysoki stopień zgodności ($Rho=0,76$). Z kolei „sędziowie” kwalifikowali każde z zachowań do jednej z trzech kategorii, którym nadano wagi: 1, 2 i 3. Jeżeli czyn był dokonywany kilkanaście razy, waga była mnożona przez dwa. W wyniku zastosowania tej procedury 6 zachowań uzyskało wagę 1, 19 zachowań wagę 2, dwa zaś zachowania wagę 3. Najwyższą wartość punktową przyznali Christie i in. stosunkowi seksualnemu z dziewczyną, która nie ukończyła 14 roku życia i rozbojowi. Natomiast już stosunek seksualny z dziewczyną w wieku 14—16 lat miał o jeden punkt mniej. Można żywić jednak pewne obawy co do porównywalności tych czynów. Poza tym nie wydaje się mimo wszystko, by społeczne niebezpieczeństwo rozboju było równe stosunkowi seksualnemu z prawidłowo rozwiniętą czternastoletnią dziewczyną. Jest to jednak generalny problem procedury ważenia: usiłuje się sprowadzić do wspólnego mianownika ze swej istoty nieporównywalne rodzaje zachowań.

W badaniach Hindelanga⁸² wagę czynów określało pięciu specjalistów na skali pięciopunktowej. Autor ten ustalał ponadto, jak sami badani oceniają dane zachowania (na skali typu likretowskiego). Podobnie jak w badaniu Christiego i in., także w badaniach Hindelanga zachodził między sędziami dość duży stopień zgodności w ocenach wagi danego zachowania. Ujawniła się ponadto duża negatywna zależność pomiędzy wagą danego czynu a częstością jego dokonywania ($Rho=0,72$), pozytywna zaś między przeciętną wagą czynu a rozmiarami jego dezaprobaty ($Rho=0,82$).

Nieco inną metodę ważenia zachowań obrał Hepburn⁸³. Punktem wyjścia do oceny wagi zachowania była jego przeciętna ocena dokonana przez samych badanych. W procedurze tej oskarżeni są jak gdyby także i sędziami własnych zachowań. W porównaniu z uprzednio analizowanymi metoda ta ma jeszcze więcej mankamentów. Wiadomo bowiem skądinąd, że zachodzi wysoka korelacja między stopniem potępienia danego zachowania a częstością jego dopuszczania się⁸⁴. Zatem ustalanie wagi czynów na podstawie ocen dokonywanych przez osoby z tej samej zbiorowości, w której mierzy się jednocześnie częstość ich występowania, jest w istocie odmiennym sposobem pomiaru tego samego parametru.

W badaniu Ostrihanskiej i Wójcik⁸⁵ o wartości punktowej zachowań

⁸¹ N. Christie i in.: *op. cit.*

⁸² M. Hindelang: *op. cit.* (1971, 1977).

⁸³ J. Hepburn, *op. cit.*

⁸⁴ Por. M. Hindelang: *op. cit.* (1971).

⁸⁵ Z. Ostrihanska: *op. cit.*; Z. Ostrihanska, D. Wójcik: *op. cit.*

także decydowali sędziowie kompetentni. W wyniku zastosowania tej procedury nieoddanie reszty z zakupów zostało uznane za „błahe lub wątpliwe”, „nieco większą” zaś wartość punktową nadano kradzieży pieniędzy rodzicom lub komuś z rodziny oraz kradzieży roweru w celu przejechania się.

Omawiane procedury ważenia zachowań nasuwają poważne obiekcje. Wydaje się, że ewentualne straty, jakie mogą wynikać z procedury ważenia (m.in. duże prawdopodobieństwo zniekształcania faktycznych rezultatów uzyskiwanych przez badanych na skalach *self-report*), są większe niż hipotetyczne korzyści, jakie płyną z jej stosowania (większa precyzja w określaniu stopnia dewiacyjnego zaangażowania).

Zresztą pewne ustalenia empiryczne wskazują na to, że procedura ważenia nie przyczynia się do istotnego wzrostu dokładności pomiaru. Na przykład Farrington ⁸⁶ porównywał trzy rodzaje skal dewiacji: różnorodności (tj. skali, która powstaje przez proste zliczanie odpowiedzi pozytywnych), częstości (tj. skali, która powstaje w rezultacie zliczania wartości punktowych przypisanych każdej poszczególnej częstości danego zachowania) oraz skali wagi czynów (która obejmuje różnorodność, częstość i ważoną „częstość” zachowań). Okazało się, po pierwsze, że skale te są ze sobą wysoko skorelowane, po drugie zaś, że predykcyjna wartość pozornie bardziej precyzyjnego narzędzia pomiaru, jakim jest skala wagi czynów, nie była w istocie większa od — zdawałoby się — tak prymitywnego narzędzia, jakim jest skala różnorodności. Wartość predykcyjna skali różnorodności względem przyszłej dewiacji badanych wynosiła $r_{bis}=0,38$, skali częstości — $r_{bis}=0,39$, skali wagi czynów — $r_{bis} = 0,39$, natomiast dla skali uwzględniającej wszystkie trzy miary (tj. różnorodność, częstość i wagę) — $r_{bis} = 0,40$.

Wydaje się zatem, że problem różnej wagi zachowań dewiacyjnych jest w praktyce badawczej znacznie mniejszy, niż się teoretycznie zakłada. Ponieważ skale ważne nasuwają wiele wątpliwości tak merytorycznych, jak i statystycznych, właściwsze wydaje się stosowanie prostszych typów skal — np. różnorodności czy częstości⁸⁷.

8. SPOSOBY UZYSKIWANIA INFORMACJI

Do tej pory mówiliśmy wyłącznie o anonimowych kwestionariuszach typu *self-report*. Stąd mogło powstać mylne wrażenie, że jest to jedyny sposób uzyskiwania informacji o rozmiarach nie ujawnionej przestępczości i innych zachowań dewiacyjnych. Tymczasem można wskazać co najmniej kilka alternatywnych bądź posiłkowych metod.

⁸⁶ D. Farrington: op. cit. (1973).

⁸⁷ Dalszego potwierdzenia tezy, że skale ważne nie są lepszymi predyktorami dewiacji w porównaniu ze skalami różnorodności i częstości, dostarczyły m.in. także badania Shapland (op. cit.) i Morash (op. cit.).

Przede wszystkim jednak rozmaite mogą być techniki prowadzenia badania za pomocą anonimowego kwestionariusza. Najczęściej badania tego typu przeprowadza się w szkołach, co jest o tyle zrozumiałe, że jest to sposób najmniej kosztowny i czasochłonny. Stosunkowo niewiele wiadomo jednak o innych sposobach badania ankietą anonimową.

Do nielicznych wyjątków należy badanie Skirbekka⁸⁸, które zmierzało do ustalenia różnic w stopniu szczerości wypowiedzi w zależności od sytuacji badawczej. Autor ten porównywał rezultaty, jakie uzyskano na losowej próbie młodzieży z terenu Oslo w zależności od zastosowanej techniki przeprowadzania badania. W pierwszej grupie przeprowadzono wywiady w domu badanych, druga grupa wypełniała anonimowy kwestionariusz, który następnie zwracała przedstawicielom ekipy badawczej, trzecia grupa także wypełniała anonimowy kwestionariusz, lecz zwracała go pocztą, czwarta grupa zaś zarówno otrzymywała, jak i zwracała kwestionariusz listownie. Wprawdzie stosunkowo niski odsetek zwrotów czynił wyniki niezbyt klarownymi, to jednak okazało się, że można mówić o tendencji do częstego przyznawania się do zachowań dewiacyjnych wraz ze wzrostem stopnia anonimowości.

Z kolei jednak np. Krohn, Waldo i Chiracos⁸⁹ nie stwierdzili statystycznie istotnych różnic między wywiadem a ankietą anonimową, jeśli chodzi o odsetki respondentów przyznających się do danego typu zachowania⁹⁰, aczkolwiek pewne różnice na niekorzyść wywiadu jednak się zarysowały.

Ankieta anonimowa ma wiele zalet. Jest w oczach respondentów uważana za bardziej „bezpieczną”, co zwiększa prawdopodobieństwo szerszych odpowiedzi, jest niekłopotliwa w stosowaniu i tania. Ma jednak także i dość poważne wady. Przede wszystkim przy zastosowaniu ściśle anonimowej ankiety nie możemy porównywać uzyskanych rezultatów z innymi źródłami informacji, ponieważ trudno jest zidentyfikować poszczególne osoby⁹¹. Stąd alternatywną metodą jest ankieta imienna. Stosując ten rodzaj ankiety *self-report* możemy dokonywać porównań z zewnętrznymi źródłami informacji i identyfikować poszczególnych respondentów, co może mieć duże znaczenie w ewentualnym ustaleniu rzetelności i trafności zastosowanego

⁸⁸ S. Skirbekk: *On Measurement of Self-reported Crimes*, 1962, tekst powielony.

⁸⁹ M. Krohni in.: *op. cit.*

⁹⁰ Podobne badania przeprowadzili także S. Shokam, N. Gera, R. Markowski i N. Kaplinsky: *Internalization of Norm. Risk Perception and Anxiety as Related to Driving Offences*, „The British Journal of Criminology”, 1976, vol. 16. Porównywali oni imienne i anonimowe kwestionariusze typu *self-report* do badania rozmiarów przestępstw i wykroczeń drogowych. Analogicznie jak Krohn i in. (*op. cit.*) także i ci autorzy nie stwierdzili różnic między kwestionariuszem imiennym i anonimowym pod względem szczerości wypowiedzi respondentów.

⁹¹ Choć nie jest to niemożliwe. Niekiedy dokonuje się takiej identyfikacji na podstawie danych biograficznych badanych — np. roku i miejsca urodzenia. Por. R. Hardt, S. Peterson-Hardt: *op. cit.*

narzędzia pomiaru. Z kolei jednak wzrasta tu prawdopodobieństwo, że respondenci będą obawiali się konsekwencji przyznania się, w związku z czym ich odpowiedzi będą mniej szczerze. Orędownikiem ankiety imiennej jest Hirschi⁹². Jest on zdania, że między ankietą imienną a anonimową brak jest istotnych różnic, jeśli chodzi o szczerość wypowiedzi respondentów. Znalazło to potwierdzenie w prowadzonych przez niego badaniach. Dlatego też kwestionariusz opracowany przez Hirschiego nie jest anonimowy.

Jednak zarówno ankieta imienna, jak i anonimowa mają wady charakteryzujące generalnie to narzędzie badawcze (m.in. trudności interpretacyjne, omyłkowe zrozumienie treści pytania, świadome lub nieświadome przeoczenia bądź pominięcia pewnych pytań etc.). Dlatego stosuje się też inne metody.

Belson⁹³ opracował i wykorzystywał z dużym powodzeniem metodę, która polega na sortowaniu kart z wyszczególnionymi typami zachowań dewiacyjnych w trakcie zaprogramowanego wywiadu. Wstępnie respondenci sortują karty na dwie sterty w zależności od typu odpowiedzi — „tak” lub „nie”. W drugim etapie badania następuje ponowne przeglądanie kart typu „nie” w celu ewentualnego przypomnienia sobie, czy dane zdarzenie istotnie nie zaszło. Z kolei (w zależności od potrzeb i celów badania) dochodzi do uściślenia sytuacji, w których nastąpiły dane czyny lub(i) częstość ich dokonywania. Warto podkreślić, że we wstępnej fazie badania (tj. podczas pierwszego sortowania kart) zagwarantowane jest jednak pewne minimum intymności, ponieważ odbywa się to niewerbalnie. Ponadto respondenci uzyskują od osób prowadzących wywiady wszystkie inne rutynowe zapewnienia o całkowitej dyskrecji. Ta metoda uzyskiwania danych o przestępczości nie ujawnionej jest szczególnie rozpowszechniona w kręgu brytyjskich badaczy zjawiska dewiacji⁹⁴.

Stosuje się też mniej zobiiektywizowane i precyzyjne metody. Short i Strodtbek⁹⁵ np. ustalali stopień zaangażowania dewiacyjnego badanych w trakcie wywiadów przeprowadzanych w czasie trwania obserwacji uczestniczącej.

Źródłem informacji bywają też osoby w rodzaju pracowników społecznych, nauczycieli czy kolegów badanych⁹⁶. Jednak dane tego typu są najczęściej bardzo ogólne i nieprecyzyjne, dlatego też mogą być użyteczne w sytuacjach, w których chodzi o przybliżone ustalenie rozmiarów prze-

⁹² T. Hirschi: *op. cit.*

⁹³ W. Belson: *op. cit.* (1968, 1976).

⁹⁴ Por. np. H. Gibson: *op. cit.*; H. Gibson, S. Morison, D. West: *op. cit.*; D. Farrington: *op. cit.* (1977); J. Shapland: *op. cit.*

⁹⁵ J.F. Short, F. Strodtbeck: *Group Process and Gang Delinquency*, Chicago 1965.

⁹⁶ Warto dodać, że analogiczne sposoby są także wykorzystywane do określania prawdomówności badanych.

stępczości badanych. Natomiast metoda ta jest mniej przydatna, gdy mierzymy do dokładnego oszacowania rozmiarów dewiacji każdego badanego.

Ostatnio coraz częściej stosuje się metodę planowego wywiadu. Metoda ta, którą m.in. zastosował w swych badaniach Gold⁹⁷, była także stosowana później m.in. przez Sheplanda⁹⁸ (w połączeniu z sortowaniem kart), Wilsona⁹⁹ oraz Morasha¹⁰⁰. Jej niewątpliwa zaleta polega na tym, że można „drażnić” respondenta tak długo, aż wydobędzie się zeń pożądane informacje. Metoda planowego wywiadu pomaga respondentom w przypominaniu sobie określonych zachowań. Umożliwia także wychwytywanie sprzeczności oraz tendencji do skrywania lub przesady. Jest ona jednak bardzo czasochłonna, stąd może być wykorzystywana do stosunkowo nie wielkich prób. Ponadto w trakcie wywiadu zachodzić mogą nieprzewidziane interakcje między osobą prowadzącą wywiad a respondentem, których nie można kontrolować w tej metodzie uzyskiwania informacji.

Przykładowo: respondentowi może spodobać się ankier, w związku z czym będzie starał się udzielić maksymalnej liczby pozytywnych odpowiedzi na pytania o zachowania dewiacyjne (jeżeli dojdzie do wniosku, że tego właśnie się od niego oczekuje). Badany może także świadomie skrywać prawdę przed ankierem, jeśli przeprowadzanie wywiadu odbywać się będzie w nieprzyjemnej atmosferze. Tego rodzaju nieprzewidzianych okoliczności może być oczywiście znacznie więcej.

Brak jest na razie dokładniejszych danych na temat różnic w poziomie ujawnianych zachowań, jakie występują przy stosowaniu metody wywiadu planowego w porównaniu z kwestionariuszem ściśle anonimowym. Jednak nieporównywalnie większa efektywność badań, w których wykorzystuje się ankietę anonimową, sprawia, że metoda ta ma znacznie więcej zwolenników w porównaniu z wywiadem.

9. RZETELNOŚĆ I TRAFNOŚĆ (UWAGI OGÓLNE)

Główna wątpliwość, jaka bywa wysuwana pod adresem kwestionariuszy typu *self-report*, dotyczy prawdomówności respondentów¹⁰¹. Generalnie rzecz ujmując, problem szczerości wypowiedzi czy szerzej — kwestia rzetelności i trafności — rysuje się w badaniach tego typu podobnie jak we wszystkich innych badaniach, w których wykorzystuje się ankietę anonimową. Nie jest to więc zagadnienie specyficzne akurat dla badań przestępczości nie ujawnionej czy innych przejawów nieprzystosowania społecznego. Wiele wskazuje na to, że kwestionariusze typu *self-report* nie prezentują się pod tym względem ani lepiej, ani gorzej w porównaniu z in-

⁹⁷ M. Gold: op. cit (1970).

⁹⁸ J. Shepland: op. cit.

⁹⁹ H. Wilson: op. cit.

¹⁰⁰ M. Morash: op. cit.

¹⁰¹ M.in. I. Deutscher: op.

nymi kwestionariuszami tego rodzaju. Jednak wielu badaczom kwestie rzetelności i trafności kwestionariuszy do badania przestępczości nie ujawnionej rysują się znacznie ostrzej niż np. kwestionariuszy do badania zachowań seksualnych, który to temat jest bodaj jeszcze bardziej drażliwy. Wynika to z faktu, iż zasadniczym celem badań *self-report* jest skłonienie respondentów do ujawnienia postępowania sprzecznego z zasadami współżycia społecznego, najczęściej zgoła przestępczego. Istnieje więc całkiem realna groźba, że badani będą rozmyślnie zatajali czyny sprzeczne z prawem w obawie przed konsekwencjami przyznania się, pomimo że większość kwestionariuszy *self-report* ma charakter anonimowy. Niebezpieczeństwo skrywania prawdy jest jeszcze większe w tych badaniach *self-report*, w których stosuje się kwestionariusze imienne. Aczkolwiek problem prawdomówności respondentów nie rysuje się obecnie już tak ostro jak w początkach stosowania metody typu *self-report*, to jednak w dalszym ciągu ogniskuje wiele emocji. Tym bardziej że istotnie metodologia badań typu *self-report* nie uległa jakimś radykalnym zmianom na lepsze w porównaniu z pionierskimi pracami Shorta i Nye'a.

Wiele racji ma Farrington¹⁰² twierdząc, że znacznemu ilościowemu rozwojowi badań typu *self-report* nie towarzyszył bynajmniej równie poważny wzrost ich technicznego i metodologicznego zaawansowania. Na pewnym etapie rozwoju badań techniką *self-report* uznano, że rzetelność i trafność tego narzędzia nie budzi zastrzeżeń, w związku z czym sfera metodologiczna tych badań jest obecnie w zastoju — konkluduje Farrington.

10. RZETELNOŚĆ

O ile zagadnieniu trafności metody *self-report* poświęcono bardzo wiele uwagi, to relatywnie mało badań zmierzało do ustalenia rzetelności tego narzędzia pomiaru, którą na użytek niniejszego opracowania rozumieć będziemy jako stałość wyników w czasie¹⁰³.

Zanalizujemy najpierw kwestię rzetelności, bo — jak słusznie podnosi Smart¹⁰⁴ — kwestionariusz nierzetelny nie może być zarazem trafny. Jednak szczupłość materiału empirycznego dotyczącego rzetelności kwestionariuszy typu *self-report* nie dziwi, jeśli zważyć bardzo poważne prze-

¹⁰² D. Farrington: *op. cit.* (1973).

¹⁰³ Pojęcie rzetelności i trafności miewają różne konotacje. Na przykład D. Farrington (*op. cit.*, 1973) nawet w stosunku do stałości wyników w czasie używa określenia „trafność predykcyjna”, nie zaś „rzetelność”. Z kolei H. Hardt i S. Peterson-Hardt (*op. cit.*) w ogóle nie posługują się pojęciem rzetelności zastępując je luźnym sformułowaniem „jakość badań *self-report*”. W dalszych rozważaniach będziemy używali tych pojęć w znaczeniu najbardziej rozpowszechnionym. Zgodnie z tym rozumieniem (np. R. Smart: *op. cit.*) kwestionariusz rzetelny to — najogólniej mówiąc — taki, którego wyniki są w pewnych granicach stałe, niezależnie od analizowanej populacji, trafny zaś to taki, który mierzy to, co ma mierzyć.

¹⁰⁴ R. Smart: *op. cit.*

szkody techniczne (a w wielu krajach także prawne), które napotyka badacz usiłujący dokonać retestu na tej samej próbie lub choćby jej części. Pomijając nawet trudności organizacyjne, możliwość zbadania stałości wyników w czasie jest najczęściej limitowana anonimowym charakterem ankiety. O ile bowiem ankieta ma charakter całkowicie anonimowy (a nie pozornie anonimowy, tj. taki, w którym w istocie dokonuje się indentyfikacji poszczególnych osób), nie ma praktycznie możliwości porównywania rezultatów uzyskiwanych przez poszczególnych badanych w teście i w re-teście. Wówczas można jedynie porównywać dane zagregowane, dotyczące pewnej większej jednostki analizy (np. porównywanie rozkładów wyników w tej samej klasie). Przeprowadzenie retestu utrudniają ponadto jeszcze dwie inne okoliczności.

Po pierwsze, charakter badanej zmiennej sprawia, że trzeba się liczyć z możliwością zmian w czasie, które wcale nie muszą oznaczać, że narzędzie pomiaru jest nierzetelne. Po prostu część badanych mogła się dopuszczać pewnych czynów w okresie między testem a retestem (jest to szczególnie prawdopodobne w badaniach, w których stosuje się kwestionariusze z przewagą pytań o stosunkowo błahe czyny). Tak więc wyniki retestu mogą być wyższe od wyników testu ¹⁰⁵, nie powinny być natomiast niższe. W tej sytuacji można żywić już uzasadnione podejrzenia co do rzetelności zastosowanego narzędzia pomiaru. Gdy bowiem wyniki retestu są niższe od wyników testu, oznacza to, że bądź w teście, bądź w re-teście część respondentów udzieliła świadomie nieprawdziwych odpowiedzi (przejaskrawiła lub ukryła fakty) lub też zawiodła ich pamięć.

Po drugie, gdy jednostką analizy jest, przypuśćmy, klasa, możemy porównać jedynie globalne rozkłady. Natomiast nie jesteśmy w stanie oszacować, w jakiej mierze pozornie analogiczny profil rozkładu jest wynikiem zmian w jego obrębie.

Założmy, że na pytanie testu o pobicie dla zwykłej fantazji 90% badanych w klasie A udzieliło odpowiedzi negatywnej, 10% badanych zaś odpowiedzi pozytywnej. Założmy ponadto, że profil rozkładu retestu był identyczny. Oczywiście, test χ^2 nie wykaże różnic w tych rozkładach. Mimo to jednak pozornie analogiczny profil rozkładu retestu może być wynikiem drastycznych zmian w jego obrębie: np. wszyscy respondenci, którzy na pytanie testu udzielili negatywnej odpowiedzi, w re-teście mogli udzielić odpowiedzi pozytywnej. Oczywiście, takie sytuacje są w praktyce bardziej niż sporadyczne, niemniej i z tym problemem należy się liczyć.

Sygnalizowane trudności sprawiają, że problem stałości w czasie wyników badań typu *self-report* bywa rzadko podejmowany. Poniżej przytaczamy ich wyniki.

¹⁰⁵ Zależy to również od czasu, jaki upłynął między testem a retestem: czym większa odległość czasowa między testem a retestem, tym większe prawdopodobieństwo, że wyniki retestu będą wyższe.

Badanie Dentlera i Monroe'go¹⁰⁶ wykazało, że na każde z pięciu pytań o zachowania dewiacyjne (kradzieże) co najmniej 92% respondentów udzieliło w reteście takiej samej odpowiedzi. Jednakże retest przeprowadzony został w dwa tygodnie po teście, a więc trochę za wcześnie, ponieważ badani mogli jeszcze pamiętać swoje odpowiedzi w teście. Belson¹⁰⁷ z kolei dokonał retestu w jeszcze krótszym terminie w stosunku do testu — w ciągu jednego tygodnia. Przeciętny odsetek zgodności wyniósł w tym badaniu 88. Farrington¹⁰⁸ natomiast przyjął, jak się wydaje, zbyt dużą cezurę czasową — dwa lata. Test przeprowadzony został na próbie chłopców w wieku 14—15 lat, retest zaś przeprowadzono na tej samej próbie, kiedy chłopcy mieli 16—17 lat. Przez ten czas mógł ulec zmianie nie tylko poziom dewiacji (co zrozumiałe), ale także sposób interpretacji i rozumienia pytań, sposób postrzegania własnych zachowań etc. Nie sposób nie wspomnieć wreszcie i o czynniku pamięci. Niezależnie od tendencji do szczerych wypowiedzi okres dwóch lat jest wystarczająco długi, aby zapomnieć wiele szczególnie błahych czynów, w kwestionariuszu zaś, którym posługiwał się Farrington, większość pytań dotyczyła właśnie stosunkowo mniej poważnych zachowań dewiacyjnych. Stąd rezultaty uzyskane przez Farringtona należy interpretować z dużą ostrożnością. Niemniej mimo rzadko spotykanej w badaniach tego typu cezury czasowej między testem i retestem wyniki uzyskane przez Farringtona świadczą o nadspodziewanej rzetelności kwestionariuszy typu *self-report*; odsetek jawnie sprzecznych odpowiedzi w stosunku do wszystkich odpowiedzi wyniósł 6,4. Jeśli podzielimy to przez dwa (tj. proporcjonalnie błędy testu do błędów retestu), okazuje się, że odsetek błędów w teście lub w reteście wynosi około 3,2.

Ponieważ badanie nie było anonimowe, można było ustalić, jakie czynniki wpływają na brak konsekwencji w sposobie udzielania odpowiedzi. Farringtona interesowały w szczególności czynniki, które warunkują sprzeczności w odpowiedziach na pytania dotyczące poważniejszych czynów tych respondentów, którzy przyznali się do nich w teście, natomiast zaprzeczyli w reteście. Okazało się, że przestępcy „ujawnieni” oraz badani, którzy uzyskali wysokie wyniki na skalach dewiacji, mieli mniejszą tendencję do zaprzeczania w reteście, że dokonali danych czynów, w porównaniu z chłopcami o niskich wynikach na skalach dewiacji i nie figurujących w policyjnych rejestrach. Tendencja do zaprzeczania w reteście była ponadto związana z wynikami uzyskiwanymi na skali kłamstwa: chłopcy o wysokich wynikach na skali kłamstwa wykazywali istotnie większą od oczekiwanej tendencję do udzielania w reteście negatywnych odpowiedzi na pytania, na które w teście udzielili odpowiedzi pozytywnej.

Generalnie rzecz biorąc tendencja do sprzecznych odpowiedzi była

¹⁰⁰ R. Dentler, L. Monroe: *op. cit.*

¹⁰⁷ W. Belson: *op. cit.* (1968).

¹⁰⁸ D. Farrington: *op. cit.* (1973).

znacznie mniej wyraźna, gdy za jednostkę analizy przyjęto nie stosunek potwierdzeń do zaprzeczeń na poszczególne pytania, lecz całościowe wyniki na skali dewiacji. Gdy podzielono rezultaty chłopców z testu na kwartyle odpowiadające poziomowi dewiacji, okazało się, że jedynie 45 spośród 405 badanych znalazło się w reteście w niższym niż uprzednio kwartylu (11,5%). Potwierdza to przypuszczenie, że dane zagregowane (np. w postaci wyników na skali dewiacji) stanowią lepszą podstawę do analiz w porównaniu z poszczególnymi pytaniami, co już uprzednio sygnalizowano.

Analogiczną procedurę określania stałości w czasie wyników kwestionariusza *self-report* zastosowała Shapland¹⁰⁹. Porównywała ona wyniki uzyskiwane przez respondentów w wieku 11—12 lat oraz w wieku 13—14 lat. Wprawdzie ogólny odsetek potwierdzeń, które przerodziły się w zaprzeczenia, był wysoki (25,67%), to jednak w wyniku dokonania bardziej pogłębionej analizy stałość wyników w czasie okazała się spora. Porównanie (jawnie) sprzecznych odpowiedzi ze wszystkimi odpowiedziami wykazało, że stosunek odpowiedzi sprzecznych do wszystkich odpowiedzi jest bardzo niski (9,66%) i to mimo dwóch lat, jakie upłynęły od pierwszego badania. Zdaniem Shapland taka metoda określania stałości wyników w czasie może być stosowana szczególnie wówczas, gdy chodzi o ustalenie generalnej tendencji do zatajania, przejawiania bądź zapominania, ponieważ dokonuje się tu jak gdyby szacunku globalnego, określa się poziom stałości w odniesieniu do wszystkich odpowiedzi, a nie tylko porównań odpowiedzi na poszczególne pytania.

Autorka czyniła jednak jeszcze bardziej ogólne porównanie stałości w czasie wyników badania *self-report*. Respondenci 11—12-letni zostali podzieleni na cztery grupy odpowiadające różnemu poziomowi dewiacji. Następnie porównano, jak przedstawiał się ich poziom dewiacji w wieku 13—14 lat. Okazało się, że jedynie dwóch z 51 badanych miało w drugim etapie badania wynik niższy niż osiągnięty w etapie pierwszym.

Haberman i in.¹¹⁰ badali stałość wyników w czasie kwestionariusza do badania narkomanii. Retest przeprowadzony w dwa tygodnie po właściwym badaniu wykazał jedynie 4% niezgodności odpowiedzi w porównaniu do rezultatów testu.

Wprawdzie niewielka liczba badań poświęconych zagadnieniu stałości w czasie kwestionariuszy *self-report* nie upoważnia do formułowania zbyt daleko idących wniosków, wydaje się jednak, że zarówno cytowane powyżej, jak i inne badania wskazują na ich dużą rzetelność.

¹⁰⁹ J. Shapland: *op. cit.*

¹¹⁰ Cyt. za R. Smart: *op. cit.*

II. TRAFNOŚĆ

Jeżeli ustalili się, że dane narzędzie pomiaru jest rzetelne, można przypuszczać, że jest ono także trafne. Wydaje się bowiem mało prawdopodobne, byśmy mieli tu do czynienia z systematycznym błędem pomiaru (choć oczywiście i takiej ewentualności wykluczyć nie można). W przeciwieństwie do badań poświęconych kwestii rzetelności kwestionariusza do badania nie ujawnionych zachowań dewiacyjnych badań, w których zajmowano się trafnością tego narzędzia, jest bardzo dużo. Dlatego też w poniższych rozważaniach będziemy przytaczali jedynie przykłady stosowanych metod i wybranych ustaleń empirycznych.

Hardt i Peterson-Hardt¹¹¹ wyróżniają następujące metody szacowania trafności kwestionariuszy *self-report*:

1) porównywanie informacji uzyskiwanych za pomocą kwestionariusza *self-report* z innymi, „zewnętrznymi” źródłami informacji, których wiarygodność nie budzi zastrzeżeń;

2) porównywanie wyników uzyskiwanych metodą *self-report* z poziomem dewiacji w zbiorowościach, których cechy zostały już wcześniej precyzyjnie określone (jest to tzw. „znana grupa” — np. młodzież z zakładów karnych bądź poprawczych);

3) porównanie wyników kwestionariuszy *self-report* ze skalami kłamstwa lub skalami do mierzenia tendencji do przedstawiania się w korzystnym świetle;

4) ustalanie trafności treściowej.

Prezentację wyników badań będziemy prowadzili zgodnie z powyższym schematem.

Dość rozpowszechnionym sposobem ustalania trafności narzędzi do badania nie ujawnionych zachowań dewiacyjnych jest porównywanie odpowiedzi badanych na pytania o zachowania dewiacyjne z innymi wiarygodnymi źródłami informacji dotyczącymi tych zachowań. Jest to jednak niestety możliwe jedynie w przypadkach kwestionariuszy imiennych lub pozornie anonimowych. Potwierdzenia prawdomówności respondentów szuka się m.in. w rejestrach policyjnych, w opiniach nauczycieli, pracowników społecznych lub nawet u kolegów badanych. Można także pytać o prawdomówność i samych zainteresowanych, którzy w innej sytuacji badawczej (wywiady przeprowadzane w jakiś czas po badaniu ankietą) stanowią jak gdyby sami dla siebie zewnętrzne źródło informacji.

I tak np. Erickson i Empey¹¹² ustalili w wyniku przeprowadzonych później wywiadów, że żaden z badanych nie zataił faktu swego kontaktu

¹¹¹ R. Hardt, S. Peterson-Hardt: op. cit.

¹¹² M. Erickson, R. Empey: *Court Record, Undetected Delinquency and Decision-making*, „The Journal of Criminal Law, Criminology and Police Science”, 1963, vol. 54.

z organami ścigania czy wymiaru sprawiedliwości ani też przestępstw, o których dokonanie był oskarżony bądź skazany.

Najczęściej jednak porównuje się wyniki kwestionariuszy *self-report* z danymi z rejestrów policyjnych lub sądowych.

I tak np. Gold¹¹³ stwierdził, że prawdopodobieństwo kontaktu z wymiarem sprawiedliwości spada monotonicznie wraz ze spadkiem częstości przestępstw ujawnionych w ankiecie *self-report*. Gibson, Morison i West¹¹⁴ stwierdzili duży stopień zgodności (83,3%) pomiędzy przestępstwami ujawnionymi za pomocą metody *self-report* a danymi na temat tej przestępczości, które zawierały rejestry policyjne. Autorzy stawiają jednak hipotezę, że sprawcy ujawnionych już czynów mogli mieć mniejsze zahamowania przed przyznaniem się do ich popełnienia¹¹⁵. Hindelang¹¹⁶ ustalił, że istnieje wyraźna, pozytywna zależność między wysokim wynikiem na skali dewiacji a figurowaniem w rejestrach policyjnych¹¹⁷.

Nieco odmiennie podchodzili do zagadnienia West i Farrington¹¹⁸ oraz Farrington¹¹⁹. Autorzy ci zajmowali się zagadnieniem tzw. trafności predykcyjnej kwestionariusza *self-report*. Dążyli oni do ustalenia, w jakim stopniu wyniki respondentów w czasie A są skorelowane z ich ujawnioną przestępczością w czasie B. Okazało się, że istnieje w tej mierze bardzo wyraźna zależność. Badani, którzy w okresie A uzyskiwali najwyższe wyniki na skali *self-report*, w okresie B znacznie częściej figurowali w policyjnych rejestrach. Rezultat ten jest istotny z dwóch względów. Po pierwsze, potwierdza dość dobitnie, że kwestionariusz typu *self-report* istotnie mierzy to, co ma mierzyć, tj. dewiację i przestępczość (a nie np. tendencję części respondentów do niezbyt stosownych żartów z badaczy). Po drugie, okazało się, że kwestionariusz *self-report* może być także użytecznym narzędziem w prognozowaniu przyszłej, już ujawnionej przestępczości. Analogiczne rezultaty uzyskał Gould¹²⁰. Porównywał on wyniki skali Shorta-Nye'a z wynikami skali ujawnionej przestępczości. Korelacja okazała się pozytywna, wysoka i istotna statystycznie. Istnienie tej zależności potwierdziło także badanie Shapland¹²¹.

W świetle przytoczonych rezultatów można uznać, że istnieje dość ściśle związki między wynikami, jakie uzyskują badani w kwestionariuszach

¹¹³ M. Gold: *op. cit.* (1970).

¹¹⁴ H. Gibson, S. Morison, D. West: *op. cit.*

¹¹⁵ Podobne stanowisko prezentuje także D. Farrington: *op. cit.* (1973).

¹¹⁶ M. Hindelang: *op. cit.* (1971, 1976).

¹¹⁷ W analogiczny sposób porównywali dane kwestionariuszy *self-report* z rejestrami policyjnymi również J. Blackmore (*op. cit.*); R. Hardt, S. Peterson-Hardt (*op. cit.*) i J. Shapland (*op. cit.*).

¹¹⁸ D. Farrington, D. West: *A Comparison between Early Delinquents and Young Aggressives*, „The British Journal of Criminology”, 1971, vol. 11.

¹¹⁹ D. Farrington: *op. cit.* (1973).

¹²⁰ L. Gould: *op. cit.*

¹²¹ J. Shapland: *op. cit.*

self-report, a danymi o ich rejestrowanych zachowaniach przestępczych bądź wykroczeniach, choć oczywiście poziom nie ujawnionej przestępczości jest znacznie wyższy od ujawnionej¹²². Porównywanie wyników badań *self-report* z urzędowymi danymi dotyczącymi przestępczości respondentów wskazuje, że obydwa źródła informacji odnoszą się do tego samego zjawiska.

Próbowano również porównać rezultatów uzyskiwanych przez badanych w kwestionariuszu *self-report* także z innymi, nieurzędowymi źródłami informacji, jak koledzy lub nauczyciele. Np. Jessor i in.¹²³ wykazali, że wyniki skali *self-report* pokrywają się z ocenami stopnia zaangażowania dewiacyjnego dokonywanymi przez nauczycieli i kolegów badanych.

Gold¹²⁴ porównywał wypowiedzi badanych na temat ich zachowań dewiacyjnych z ocenami tych zachowań, których dokonywali ich koledzy i nauczyciele. Zależności okazały się wyraźne¹²⁵ i zgodne z oczekiwaniami. Odminną metodę szacowania trafności kwestionariusza *self-report* zaproponowali Toby i Toby¹²⁶. Twierdzili oni, że poziom ujawnionej przestępczości kolegów badanego stanowi lepszy wskaźnik dewiacji respondentów tak w porównaniu z metodą *self-report*, jak i danymi urzędowymi. Późniejsze testy weryfikujące tę hipotezę nie wykazały bezpośredniego związku między rejestrowaną przestępczością kolegów badanych a ich własną przestępczością ujawnianą za pomocą metody *self-report*. Hardt i Peterson¹²⁷ wykazali jednak, że w pewnym ograniczonym zakresie informacje o ujawnionej przestępczości kolegów badanych mogą stanowić pośrednie potwierdzenie ich własnej przestępczości stwierdzonej za pomocą metody *self-report*.

Oryginalny sposób określania prawdopodobności respondentów przez porównanie ze źródłem zewnętrznym wykorzystali w swym badaniu Hardt i Peterson-Hardt¹²⁸. Otóż przez wiele lat władze miejskie w miejscowości, w której realizowane było badanie, miały poważny problem z plagą włamań do liczników parkingowych. Wśród młodzieży zapanowała swoista moda na okradanie tych automatów, opracowane zostały przemysłne technologie dokonywania tego przestępstwa, tak że dochody miasta z tytułu opłat parkingowych znacznie zmalały. W tej dość drastycznej sytuacji władze miejskie postanowiły zainstalować nowy, znacznie droższy, lecz bezpieczniejszy typ liczników parkingowych, co natychmiast spowodowało

¹²² Np. R. Hardt, S. Peterson - Hardt: *op. cit.*

¹²³ R. Jessor, T. Graves, R. Hanson, S. Jessor: *Society, Personality and Deviant Behavior*, New York 1968.

¹²⁴ M. Gold: *op. cit.* (1963).

¹²⁵ Por. także H. Wilson: *op. cit.*

¹²⁶ J. Toby, M. Toby: *Law School Status as Predisposing Factor in Sub-cultural Delinquency*, tekst powielony, niedatowany.

¹²⁷ R. Hardt, S. Peterson: *op. cit.*

¹²⁸ R. Hardt, S. Peterson - Hardt: *op. cit.*

znaczny spadek włamań do tych urządzeń. Wśród innych pytań o zachowania dewiacyjne autorzy zamieścili także pytanie o okradanie liczników parkingowych. Jak pamiętamy, zestaw odpowiedzi tego kwestionariusza obejmował zarówno czyny dokonywane w ciągu roku poprzedzającego badanie, jak również w okresie wcześniejszym. Okazało się, że spośród chłopców, którzy dokonywali włamań do liczników parkingowych, tylko 28% przyznało, że dopuściło się tego w ciągu roku poprzedzającego badanie (tj. już po zainstalowaniu liczników nowego typu), podczas gdy wśród chłopców, którzy przyznali się do tego, że kiedykolwiek dopuścili się włamań do automatów innego rodzaju, aż 67% dokonało tego w ciągu roku poprzedzającego badanie. Rezultat uzyskany przez Hardta i Peterson jest istotny jeszcze i z tego względu, że wskazuje, iż kwestionariusz *self-report* może być stosunkowo czułym narzędziem do badania zmian zachowań w czasie.

Porównania wyników skal dewiacji z zewnętrznymi źródłami informacji zdają się potwierdzać dużą trafność kwestionariuszy *self-report* jako narzędzia do badania nie ujawnionej przestępczości i innych zachowań dewiacyjnych.

Inną skuteczną metodą ustalania trafności kwestionariuszy do badania nie ujawnionych zachowań dewiacyjnych jest porównywanie uzyskanych rezultatów z parametrami jakiejś znanej grupy — np. przestępczej. Skoro bowiem uznaje się dane urzędowe za wiarygodne (co samo w sobie stanowi odrębne zagadnienie) i ustala się na tej podstawie czynniki związane z przestępczością, to jeśli porównanie korelatów zachowań przestępczych w grupach zidentyfikowanych i w grupie badanej za pomocą metody *self-report* wykaże wysoki stopień zgodności w tej mierze, można na tej podstawie wnioskować zasadnie także o trafności tego alternatywnego narzędzia pomiaru.

Jednak już samo porównanie wyników, jakie uzyskuje młodzież nie uznana formalnie za dewiacyjną i młodzież, która została urzędowo uznana za przestępców (np. młodzież z zakładów poprawczych), w pewnej mierze wystarcza do stwierdzenia, że zastosowane narzędzie mierzy w istocie to, co powinno mierzyć, czyli dewiacyjne zaangażowanie. Z tych przesłanek wychodzili Short i Nye¹²⁹ i dlatego porównywali wyniki swych skal wśród młodzieży ze szkół i tej, której przestępczość została urzędowo stwierdzona. Zgodnie z przewidywaniami młodzież z zakładów poprawczych uzyskiwała znacznie wyższe wyniki na skali Sharta-Nye'a w porównaniu z młodzieżą nie karaną. Oznacza to, iż skale do badania nie ujawnionej przestępczości mierzą ten sam wymiar czy sferę zachowań, co rejestry urzędowe.

Porównanie rozkładów odpowiedzi na pytania o zachowanie dewiacyjne młodzieży „nieprzestępczej” i przestępczej zawiera bardzo wiele badań prowadzonych za pomocą kwestionariusza *self-report*. M.in. Voss¹³⁰ sto-

¹²⁹ J. F. Short, I. Nye: *op. cit.* (1957).

¹³⁰ H. Voss: *op. cit.* (1963, 1964).

sował skalę Shorta-Nye'a w celu ustalenia niektórych korelatów zachowania dewiacyjnego młodzieży różnych ras i mającej rodziców o różnej pozycji społeczno-ekonomicznej. Okazało się, że na wszystkie pytania o zachowania dewiacyjne młodzież, której przestępczość ujawniono, miała znacznie wyższe odsetki odpowiedzi pozytywnych w porównaniu z młodzieżą, która nie figuruje w rejestrach policyjnych.

Oczywiście, jest to tylko pośredni i dość słaby dowód na trafność skal do badania nie ujawnionych zachowań dewiacyjnych. Przyjmuje się tu bowiem jak gdyby dowód przez negację: skoro młodzież, której przestępczość ujawniono, nie uzyskuje niższych wyników na skali dewiacji od młodzieży, która nie miała kontaktów z wymiarem sprawiedliwości, to znaczy, że metoda *self-report* jest trafna.

Bezpośredniego dowodu na poparcie tezy o trafności metody *self-report* dostarczają natomiast te badania, w których zmierzano do porównania wyników na skali *self-report* „przestępców” i „nieprzestępców” pod kątem ich związku z innymi korelatami dewiacji. W tym przypadku rozumuje się następująco. O ile zakładamy, że wzorcem dewiacyjnego zaangażowania jest miara urzędowo stwierdzonej dewiacji (tj. innymi słowy, że np. skazany przestępca na pewno jest dewiantem), a ponadto o ile przyjmujemy, że dotychczasowe ustalenia nauki co do uwarunkowań dewiacji są wiarygodne, to jeżeli w obydwu porównywanych grupach będą występowały analogiczne korelacje ze zmiennymi, co do których przyjmuje się, że mają związek z tym zjawiskiem, oznacza to, że test *self-report* jest trafny, że mierzy to, co ma mierzyć, tj. dewiację. Może zatem stanowić alternatywną bądź uzupełniającą miarę dewiacyjnego zaangażowania w stosunku do danych urzędowych.

Już Short¹³¹ stwierdził, że tak skazani przestępcy, jak i przestępcy zgodnie z kryteriami stosowanymi w badaniach typu *self-report* nie różnią się pod względem zmiennych, które za predycyjne względem zachowania dewiacyjnego uznaje teoria zróżnicowanych powiązań. Podobieństwa te odnotowano także w innych badaniach weryfikacyjnych teorii Sutherlanda¹³². Stwierdza się zgodnie, że tak „oficjalni” przestępcy, jak i „przestępcy” ujawnieni w badaniach typu *self-report* różnią się jedynie pod względem siły związku zmiennych niezależnych z zachowaniem dewiacyjnym (i to nie zawsze), nie są to natomiast grupy różniące się jakościowo.

Istotnym wyjątkiem od tej reguły, który był zresztą jedną z przyczyn tak bujnego rozwoju badań przestępczości nie ujawnionej, jest pozycja społeczno-ekonomiczna. Zarówno w teorii dewiacji, jak i w badaniach ujawnionych przestępców stwierdza się silny związek dewiacji z niską pozycją społeczno-ekonomiczną¹³³. Natomiast przytłaczająca większość ba-

¹³¹ J. F. Short: op. cit. (1957, 1958, 1960).

¹³² Zob. A. S i e m a s z k o: *Spoleczna geneza przestępczości*, Warszawa 1979.

¹³³ Z tych przesłanek zrodziła się teoria anomii Merton. Zob. R. K. Merton: *Teoria socjologiczna i struktura społeczna*, Warszawa 1982.

dań, w których wykorzystywano metodę *self-report*, bądź w ogóle nie wykazuje istnienia takiego związku, bądź zgoła związek pozytywny zamiast negatywnego, sugerowanego przez badania ujawnionych przestępców¹³⁴. Ponieważ jest to jednak tylko wyjątek od reguły, która stanowi w tym przypadku stosowalność teorii dewiacji tak do przestępców ujawnionych, jak i nie ujawnionych, możemy przyjąć, że tu akurat wiarygodność danych urzędowych może być zasadnie kwestionowana.

Wyniki badań rozmiarów nie ujawnionych zachowań dewiacyjnych zdają się świadczyć o tym, że koncentracja zjawisk patologii społecznej w niższych warstwach społeczno-ekonomicznych, na którą wskazują statystyki policyjne i sądowe, może być artefaktem wynikającym z selektywności działania organów ścigania i wymiaru sprawiedliwości. Jednak kwestia ta pozostaje nadal otwarta. W ostatnich latach obserwujemy ponowny wzrost zainteresowania badaczy problematyką związków pozycji społeczno-ekonomicznej z zachowaniem dewiacyjnym, określanym za pomocą metody *self-report*¹³⁵. Rezultaty tych badań nie są jednak jednoznaczne.

Znaną grupą nie muszą być koniecznie ujawnieni przestępcy. Można ją wyłonić także z badanej populacji, dzieląc ją pod kątem poziomu czy nasilenia dewiacji, przy czym tak wyłonią grupa „przestępców” może, choć nie musi być jednocześnie grupą, która miała kontakt z organami ścigania. Jeśli bowiem w wyniku zastosowania takiego podziału okaże się, że grupy te różnią się także zasadniczo pod względem siły związku z analizowanymi zmiennymi niezależnymi, wówczas również stanowi to dowód na trafność kwestionariuszy *self-report*, gdyż mają one siłę dyskryminującą „przestępców” od „nieprzestępców” zgodnie z przyjętymi kryteriami.

Hardt i Peterson-Hardt¹³⁶ wyłonili „znaną grupę” spośród respondentów, którzy przyznali się, że byli oficjalnie ostrzegani przez policję za złe zachowanie się. Wychodzono z założenia, że grupa badanych, która weszła już poniekąd w konflikt z prawem, a w każdym razie miała formalny kontakt z policją, będzie miała wyższy poziom dewiacji od pozostałych. Okazało się, że chłopcy, którzy byli formalnie ostrzegani przez policję dwukrotnie

¹³⁴Pierwsze badanie tego rodzaju przeprowadzili: I. Nye, J. F. Short, V. Olson: *Socio-economic Status and Delinquent Behavior*, „American Journal of Sociology”, 1958, vol. 63; zob. także m.in. Axenroth: op. cit.

¹³⁵Zob. Axenroth: op. cit. i cytowana przez autora bogata literatura na ten temat. Tribble jednak jest zdania, że brak związku między zachowaniem dewiacyjnym a pozycją społeczno-ekonomiczną, jaki ujawnia większość badań typu *self-report*, wynika ze specyficznego sposobu definiowania „przestępczości nieletnich”. Jeśli bowiem uwzględni się wyłącznie czyny, które stanowią także przestępstwa w wypadku, gdy dopuszcza się ich dorośli, to wówczas badanie typu *self-report* także może dać rezultat zgodny z przewidywaniami Mertona. Zob. S. Tribble: *Socio-economic Status and Self-reported Juvenile Delinquency*, „Canadian Journal of Criminology and Corrections”, 1972, vol. 14.

¹³⁶R. Hardt, S. Peterson-Hardt: op. cit.

lub częściej, mieli znacznie wyższy poziom dewiacji od pozostałych chłopców. Prawie 60% respondentów ostrzeganych przez policję dwukrotnie lub częściej miało wysoki poziom na skali dewiacji, w grupie chłopców ostrzeganych jednokrotnie odsetek ten wynosił 37, natomiast wśród chłopców, którzy nigdy nie byli ostrzegani przez policję, odsetek z wysokimi wynikami na skali dewiacji wynosił tylko 15. Na podkreślenie zasługuje fakt, że była to tylko jednostronna zależność, tj. chłopcy formalnie ostrzegani przez policję istotnie znacznie częściej mieli wysoki poziom dewiacji. Gdy jednak sprawdzono, jak wyglądają kontakty z policją wszystkich chłopców, których poziom dewiacji był wysoki, okazało się, że większość z nich nie miała formalnych kontaktów z policją. Świadczy to także o użyteczności badań typu *self-report* i ich większej czułości w wychwytywaniu badanych o dużym zaangażowaniu dewiacyjnym.

Hirschi¹³⁷ podzielił badanych chłopców ze szkół na podstawie wyników uzyskiwanych przez nich na skalach dewiacji na „przestępców” i „nieprzestępców”, czyli na grupy o wysokich i niskich wynikach na skali dewiacji. Okazało się, że wyodrębnione w ten sposób grupy różniły się zasadniczo pod względem siły związku z uwzględnionymi przez tego autora grupami zmiennych niezależnych. Metoda ta jest zresztą powszechnie stosowana w badaniach typu *self-report*¹³⁸. Stwierdza się konsekwentnie z jednej strony duże różnice między „przestępcami” i „nieprzestępcami” w wyniku podziału, którego podstawą są wyniki na skali dewiacji, z drugiej zaś duże podobieństwa między „przestępcami” nie ujawnionymi a przestępcami, których czyny zostały urzędowo stwierdzone. Rezultaty powyższych badań stanowią kolejne, dobitne potwierdzenie trafności kwestionariusza typu *self-report* jako narzędzia badawczego.

Jednak o trafności testu można powiedzieć niemało patrząc po prostu na rozkłady odpowiedzi na poszczególne pytania kwestionariusza. To właśnie określa się jako trafność treściową¹³⁹. Trafność kwestionariusza wpływa bowiem pośrednio z profilu rozkładu odpowiedzi. Przykładowo: patrząc na odsetki odpowiedzi twierdzących stwierdza się zazwyczaj, że maleją one wraz ze wzrostem wagi czynu. Do czynów błahych przyznaje się bardzo wysoki odsetek badanych, maleje on zaś systematycznie wraz ze wzrostem wagi i rzadkości czynu. Gdyby odpowiedzi na pytania o zachowania dewiacyjne były przypadkowe i(lub) nieszczerze, tendencja taka nie powinna występować. Zatem już sama konfiguracja odpowiedzi może być pomocna w ustalaniu trafności metody *self-report*.

W kwestionariuszach, w których zakreślane są różne cezury czasowe, okazuje się, że respondenci przyznają się do znacznie mniejszej liczby za-

¹³⁷ T. Hirschi: *op. cit.*

¹³⁸ Zob. m.in. J. F. Short: *op. cit.* (1958, 1969); J. Hepburn: *op. cit.*; R. Simonsin.: *op. cit.*

¹³⁹ H. Błałock, *Statystyka dla socjologów*, Warszawa 1975, s. 21.

chowań dokonywanych w ciągu ostatniego roku w porównaniu do zachowań, których dopuszczali się kiedykolwiek.

Dalej stwierdza się, że poziom dewiacji chłopców jest — ogólnie rzecz biorąc — wyższy od poziomu dewiacji dziewcząt. Z pozoru jest to dość banalna konstatacja, bowiem fakt ten jest powszechnie znany. Jednak podstawą owej wiedzy są przecież dane o rozmiarach i uwarunkowaniach przestępczości ujawnionej. Gdyby kwestionariusze typu *self-report* odzwierciedlały zupełnie przypadkowe odpowiedzi respondentów, nie powinny wystąpić te skądinąd dobrze znane różnice ze względu na płeć. Podobne wyniki na skalach dewiacji są wyższe dla starszej młodzieży w porównaniu z młodzieżą młodszą. Także i w tym wypadku, gdyby odpowiedzi były przypadkowe, nie powinniśmy się spodziewać, że różnice takie będą występowały.

Większość kwestionariuszy do badania nie ujawnionych zachowań dewiacyjnych zawiera także specjalne pytania-pułapki. Ich istota polega na tym, że prawdomówny respondent, który odpowiedział pozytywnie na jedno z takich pytań, powinien także udzielić pozytywnej odpowiedzi na pozostałe pytania. Przykładowo: jeżeli w kwestionariuszu znajduje się ogólne pytanie o to, czy zdarzyło się respondentowi coś ukraść, i pytania szczegółowe o konkretne kradzieże, pozytywna odpowiedź na pierwsze pytanie implikuje także pozytywne odpowiedzi na pozostałe pytania¹⁴⁰.

Hardt i Peterson-Hardt¹⁴¹ ustalali prawdomówność respondentów m.in. przez porównywanie pytań o legitymowanie i zatrzymywanie lub aresztowanie przez policję. Uzyskano jedynie 2% niezgodnych odpowiedzi, co mieści się w normalnych granicach błędu pomiaru (może np. wynikać z przeoczenia lub opaczego zrozumienia pytania).

Niezależnie od pytań-pułapek większość kwestionariuszy typu *self-report* zawiera także skale kłamstwa, które mają na celu eliminowanie z dalszych analiz respondentów, uzyskujących na tych skalach dyskwalifikujący rezultat. Jednakże skale kłamstwa nie są zbyt precyzyjnym wskaźnikiem prawdomówności. Jakkolwiek wielu badaczy zaleca ich stosowanie¹⁴², to wyniki badań nie są w tej mierze jednoznaczne. Np. Hardt i Peterson-Hardt¹⁴³ nie stwierdzili korelacji między wynikami na skalach kłamstwa i dewiacji ($\gamma=0,06$). Może to oznaczać, że badani, którzy mają tendencję do przedstawiania się w korzystnym świetle lub po prostu kłamią, nie mają wyższego (ani też niższego) od innych respondentów wyniku na

¹⁴⁰ I. Nye i J. F. Short: op. cit. (1957); I. Nye: op. cit.; N. Christie i in.: op. cit.

¹⁴¹ P. Hardt, S. Peterson-Hardt: op. cit.

¹⁴² Np. I. Nye: op. cit.; D. Farrington: op. cit. (1973).

¹⁴³ R. Hardt, S. Peterson-Hardt: op. cit. Także R. Smart (op. cit.) jest zdania, że skale kłamstwa raczej nie zdają egzaminu w kwestionariuszach typu *self-report*.

skalach dewiacji lub też że zależność między wynikami na skalach kłamstwa i dewiacji jest inna niż prostoliniowa — np. że może być to funkcja w kształcie litery „U”. Nawiasem mówiąc, stosuje się także i dość niecodziennie metody ustalania prawdomówności respondentów. Np. Clark i Tift¹⁴⁴ badali prawdomówność respondentów za pomocą wykrywacza kłamstw.

Nie wydaje się jednak, by do ustalania trafności kwestionariusza typu *self-report* były potrzebne aż tak skomplikować metody jak wykrywacz kłamstw. Bowiem to wszystko, co zostało powiedziane do tej pory, zdaje się wskazywać jednoznacznie, że jest to zarówno trafne, jak i rzetelne narzędzie pomiaru, szczególnie dla stosunkowo dużych prób i w większej mierze dla ustalenia ogólnych wyników na skalach dewiacji niż wyników, jakie uzyskują poszczególni respondenci lub rozkładów odpowiedzi na poszczególne pytania.

IV. WNIOSKI KOŃCOWE

Obecna sytuacja w dziedzinie badań kryminologicznych, w których wykorzystuje się metodę *self-report*, jest dość paradoksalna: metoda wyraźnie zdystansowała metodologię. Znikome pocieszenie stanowić tu może fakt, że analogiczne zjawisko obserwuje się także w sferze badań empirycznych innych nauk społecznych: z reguły rodzaj zastosowań danej metody badawczej nie idzie w parze z rozwojem jej metodologii. Klasyczny przykład stanowi chociażby stan metodologii wywiadu socjologicznego w porównaniu z liczbą badań, w których metoda wywiadu jest stosowana. Jednak sła ba to pociecha. Spróbujemy więc zwrócić uwagę na te kwestie metodologiczne badań typu *self-report*, które bezwzględnie wymagają dalszej penetracji.

Po pierwsze, brak w dalszym ciągu definitywnej odpowiedzi na pytanie, w jakiej mierze ściśle anonimowa ankieta *self-report* stanowi lepsze narzędzie pomiaru zachowań dewiacyjnych w porównaniu z kwestionariuszem imiennym. Jak wspominaliśmy, przeprowadzone porównania zdają się świadczyć na rzecz ankiety anonimowej, choć rezultaty badań nie są bynajmniej jednoznaczne. Sprawa ta jest istotna o tyle, że gdyby okazało się, iż imienny kwestionariusz *self-report* jest tak samo dobrym (lub tylko nieznacznie gorszym) narzędziem pomiaru, korzystniej byłoby stosować kwestionariusze sygnowane. Rozwiązanie takie ma bowiem wiele zalet, wśród których wskazać można przykładowo możliwość określania wyborów socjometrycznych respondentów, porównywanie wyników badania z zewnętrznymi źródłami informacji czy też możliwość dotarcia do rodziców respondentów w celu przeprowadzenia z nimi wywiadu. Jak wiadomo, ankie-

¹⁴⁴ J. Clark, L. Tift: *op. cit.*

ta ściśle anonimowa wyklucza możliwości tego rodzaju. Stosowanie ankiety imiennej miałoby także wpływ na doskonalenie samej metodologii badań typu *self-report*, ponieważ ułatwiałoby znacznie np. określanie stałości wyników w czasie.

Po drugie, w dalszym ciągu nie jest definitywnie rozstrzygnięta kwestia sposobu ujmowania pytań w kwestionariuszach *self-report*: czy istotnie ujęcia eufemistyczne, już ze swej natury mniej jednoznaczne, sprzyjają większej szczerości wypowiedzi w porównaniu z ujęciami bezpośrednimi? Gdyby okazało się, że problem ten jest mniej istotny, niż się to na ogół zakłada, wiele przemawiałoby jednak za formułowaniem pytań wprost. Przykładowo: „czy zdarzało ci się ukraść...”, nie zaś „czy zdarzało ci się zabrać i nie oddać...”.

Po trzecie, należałoby określić optymalną cezurę czasową, tj. taką, która z jednej strony sprzyjałaby jak najwierniejszemu odzwierciedleniu poziomu dewiacji każdego badanego, z drugiej zaś nie powodowałaby jeszcze niepotrzebnych „szumów informacyjnych” (np. zaliczenie do grupy z wysokim poziomem dewiacji respondenta, który w przeszłości postępował naganie, lecz w chwili realizacji badania już się „ustatkował”). Jak pamiętamy z uprzednich rozważań, istnieją rozmaite propozycje w kwestii cezury czasowej, nagromadzony zaś do tej pory materiał empiryczny nie daje jeszcze podstaw do uznania konkretnego rozwiązania za lepsze od pozostałych.

Oceniając generalnie stan metodologii badań typu *self-report* nie można nie przyznać racji Farringtonowi, który określił go jako zastój¹⁴⁵. BOWIEM fakt, że jest to obecnie powszechnie stosowane (można by stwierdzić, że wręcz nadużywane) narzędzie badawcze do analizy przestępczości i innych przejawów nieprzystosowania społecznego w wielu krajach, nie oznacza niestety, że wszystkie problemy metodologiczne zostały już pomyślnie rozwiązane. Tymczasem mamy do czynienia z sytuacją, w której w kryminologii i socjologii dewiacji skale *self-report* stosuje się już bezrefleksyjnie, na zasadzie zbliżonej do stosowania testu MMPI w psychologii¹⁴⁶. Ten stan rzeczy jest bardzo zły. Niezależnie bowiem od merytorycznych zalet i wad

¹⁴⁵ D. Farrington: op. cit. (1973), s. 99.

¹⁴⁶ W tym miejscu nie od rzeczy byłoby zadać pytanie, skąd właściwie wzięła się tak duża popularność metody *self-report* w badaniach nad zachowaniami dewiacyjnymi. O jej licznych zaletach merytorycznych była już mowa we wstępie. Wydaje się jednak, że zaważyła na tym również pewna okoliczność pozamerytoryczna. Mianowicie metoda *self-report* jest bardzo tania i efektywna. Stosowanie ankiety audytoryjnej umożliwia w ciągu kilku dni realizację badania na próbie kilku tysięcy osób, przeprowadzenie zaś analogicznej liczby wywiadów trwa niekiedy kilka miesięcy i jest bardzo kosztowne. Podobnie kłopotliwe i kosztowne byłoby jednak prowadzenie badań typu *self-report* np. w zakładach poprawczych. Stąd chyba nieprzypadkowo najbardziej popularny jest najtańszy i najprostszy zarazem wariant badań metodą *self-report*: badania anonimową ankietą audytoryjną prowadzone w szkołach.

testu MMPI jest to przecież znormalizowane narzędzie badawcze, a co ważniejsze, jest to test zunifikowany. Określone są jego normy w populacji generalnej, istnieją wersje MMPI dla poszczególnych krajów, z uwzględnieniem ich specyfiki obyczajowej, kulturowej etc. Natomiast w przypadku skali *self-report* sytuacja jest odmienna. Jedyłą w miarę powszechnie stosowaną skalą jest skala Shorta i Nye'a ¹⁴⁷. Jednak możliwości jej wykorzystywania ograniczają się praktycznie do państw anglosaskich, ponadto zaś nadaje się ona do zastosowania do populacji, co do których zakłada się, że cechuje je relatywnie niski poziom dewiacji. Brak zunifikowanej skali *self-report* sprawia, że stosowane są najrozmaitsze skale spośród istniejących. Najczęściej bywa jednak tak, że poszczególni badacze sami tworzą skale *self-report* na użytek swych badań. Nierzadko nie podaje się już nawet pytań tworzących takie skale ¹⁴⁸ zakładając, że zarówno rzetelność, jak i trafność stosowanego narzędzia pomiaru jest poza sporem. Z metodologicznego punktu widzenia sytuacja taka nasuwa, mówiąc ogólnie, pewne wątpliwości.

Należałoby więc dążyć do stworzenia standardowego narzędzia pomiaru nie ujawnionych zachowań dewiacyjnych, którego parametry w populacji generalnej byłyby oszacowane. Do tego jednak jeszcze daleka droga.

¹⁴⁷ Jak wspominaliśmy, w ostatnich latach coraz częściej stosuje się także skalę Hirschiego, w której znalazły się trzy pozycje z kwestionariusza Shorta i Nye'a oraz dwie pozycje z kwestionariusza Dentlera i Monroe'go. Natomiast w badaniach *self-report*, w których prowadzi się z respondentami zaprogramowane wywiady, bywa także stosowana skala Golda.

¹⁴⁸ Np. W. Chambliss, R. Nagasawa: *On the Validity of Official Statistics. A Comparative Study of White, Black and Japanese High-school Boys*, „Journal of Research in Crime and Delinquency”, 1969, vol. 6.